Organization

Conference Committee
Martin Corley, Hugh Rabagliati, Mante Nieuwland, Patrick Sturt, Martin Pickering, Andrea Martin

Sponsorship
AMLaP 2014 has been generously supported by the School of Philosophy, Psychology and Language Sciences at the University of Edinburgh.

We are grateful to John Dewar and Sons Ltd. for their support of the Whisky Tasting event.

The Early Career Researchers’ Breakfast is kindly sponsored by Elsevier.

Reviewers
Venues and attractions

AMLaP 2014’s events occur across two locations. On Wednesday September 3 the conference kicks off with a public lecture by Pim Levelt followed by a Welcome Reception. These events will take place at the central campus of the University of Edinburgh, at the Reid Music Hall and Informatics Forum, respectively.

The remainder of the scheduled programme (Thursday through Saturday) will take place at the Queen Mother Conference Centre of the Royal College of Physicians of Edinburgh. This unique venue is located in the heart of Edinburgh’s New Town. The conference dinner will take place in the 18th century vaults beneath the Old Town, at The Caves.
## Schedule

**WEDNESDAY SEPTEMBER 3**

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>17:00</td>
<td><strong>Public Lecture: Pim Levelt</strong></td>
<td>Reid Music Hall, The University of Edinburgh</td>
</tr>
<tr>
<td>18:30</td>
<td><strong>Welcome Reception</strong></td>
<td>Informatics Forum, The University of Edinburgh</td>
</tr>
<tr>
<td>Time</td>
<td>Title</td>
<td>Speaker(s)</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>09:30</td>
<td>Eliminating unpredictable linguistic variation through interaction</td>
<td>Kenny Smith, Olga Feher, Nikolaus Ritt</td>
</tr>
<tr>
<td>09:50</td>
<td>Interference during antecedent retrieval: Effects of reading skill</td>
<td>Dave Kush</td>
</tr>
<tr>
<td>10:10</td>
<td>Are there universal prosodic cues to speech structure?</td>
<td>Laurence White, Sven Mattys, Silvia Benavides-Varela, Katalin Mådy</td>
</tr>
<tr>
<td>10:30</td>
<td>The cognitive and neural basis of pragmatic processing: A case study of jokes</td>
<td>Evelina Fedorenko, Jeanne Gallée, Zuzanna Balewski</td>
</tr>
<tr>
<td>10:50</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>11:20</td>
<td>The relationship between literacy and spoken language processing</td>
<td>Naomi Havron and Inbal Arnon</td>
</tr>
<tr>
<td>11:40</td>
<td>Adjective sprouting in much-less ellipsis</td>
<td>Katy Carlson and Jesse Harris</td>
</tr>
<tr>
<td>12:00</td>
<td>Adaptation of active gap filling in coordinated sentences</td>
<td>Akira Omaki and Aaron Apple</td>
</tr>
<tr>
<td>12:20</td>
<td>Active search may override local coherence effects</td>
<td>Lauren Ackerman, Nina Kazanina, Masaya Yoshida</td>
</tr>
<tr>
<td></td>
<td>Poster session 1 with catered lunch</td>
<td></td>
</tr>
<tr>
<td>14:30</td>
<td>Jesse Snedeker, Harvard University</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>16:00</td>
<td>Toddlers exploit prosodic structure online to constrain syntactic analysis</td>
<td>Alex de Carvalho, Isabelle Dautriche, Anne Christophe</td>
</tr>
<tr>
<td>16:20</td>
<td>Verbal morphemes guide thematic role assignment in children</td>
<td>Duygu Özge, Deniz Özkan Göktürk, Berna Uzundağ, Aylin Küntay, Jesse Snedeker</td>
</tr>
<tr>
<td>16:40</td>
<td>Learning structure-dependent processing from combinatorial meaning</td>
<td>Hartmut Fitz and Franklin Chang</td>
</tr>
<tr>
<td>17:00</td>
<td>Evidence for a phase transition in learning a recursive artificial grammar</td>
<td>Whitney Tabor, Garrett Smith, Pyeong Whan Cho</td>
</tr>
<tr>
<td>18:00</td>
<td>Whisky Tasting</td>
<td>Great Hall, RCPE</td>
</tr>
<tr>
<td>Time</td>
<td>Event</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>08:30</td>
<td><strong>Breakfast Workshop for Students and Early Career Researchers</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Getting Your Paper Published</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adam Fraser, Associate Publisher, JML</td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td><strong>SNAP judgments: Small N acceptability paradigm (SNAP) for linguistic</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>judgments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kyle Mahowald, Peter Graff, Jeremy Hartman, Edward Gibson</td>
<td></td>
</tr>
<tr>
<td>09:50</td>
<td><strong>Lexical clustering in efficient language design</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isabelle Dautriche, Kyle Mahowald, Edward Gibson, Anne Christophe,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Steven Plantadosi</td>
<td></td>
</tr>
<tr>
<td>10:10</td>
<td><strong>Does sociolinguistic distance inhibit syntactic alignment? Evidence</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>from syntactic priming between and within varieties of German</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Janine Lüthi and Constanze Vorwerg</td>
<td></td>
</tr>
<tr>
<td>10:30</td>
<td><strong>Lexical (mis)alignment: Automatic or strategic behavior?</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anouschka Foltz, Judith Gaspers, Kristina Thiele, Philipp Cimiano,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prisca Stenneken</td>
<td></td>
</tr>
<tr>
<td>10:50</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>11:20</td>
<td><strong>Input variability and the specificity of perceptual recalibration</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eva Reinisch and Holger Mitterer</td>
<td></td>
</tr>
<tr>
<td>11:40</td>
<td><strong>Lexically driven perceptual learning of systemic vowel variation: Evidence for cross-talker generalization following single-talker exposure</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kodi Weatherholtz</td>
<td></td>
</tr>
<tr>
<td>12:00</td>
<td><strong>The role of speech production in phonological decoding during visual word recognition: The case of phonotactic repair</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yue Sun and Sharon Peperkamp</td>
<td></td>
</tr>
<tr>
<td>12:20</td>
<td><strong>No place for /h/: ERP investigation of English fricative place features</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kevin Schluter, Stephen Politzer-Ahles, Diogo Almeida</td>
<td></td>
</tr>
<tr>
<td>12:40</td>
<td>Poster session 2 with catered lunch</td>
<td></td>
</tr>
<tr>
<td>14:30</td>
<td><strong>Rob Hartsuiker, University of Ghent</strong></td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>16:00</td>
<td><strong>Panel Discussion</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jesse Snedeker, Rob Hartsuiker, David Poeppel</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chair: Chuck Clifton</td>
<td></td>
</tr>
<tr>
<td>18:00</td>
<td>Conference Dinner</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Caves, Niddry Street South</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>David Poeppel, New York University</td>
<td></td>
</tr>
<tr>
<td>10:30</td>
<td>Coffee</td>
<td></td>
</tr>
</tbody>
</table>
| 11:00  | **Ambiguity advantages in attachment and pronominal reference: Evidence from eye movements during reading**  
Margaret Grant, Brian W. Dillon, Shayne Sloggett  
**Context-dependent mapping of linguistic and color representations challenges strong forms of embodiment**  
Falk Huetig and Ernesto Guerra  
**The consequences of mislocated fixations during reading**  
Laura Wakeford and Wayne Murray  
**Native and non-native processing of Italian subject pronouns: Evidence from eye-tracking**  
Tihana Kras, Patrick Sturt, Antonella Sorace |
| 12:00  | **Poster session 3 with catered lunch**                                  |
| 14:10  | **How do we compute predictions for an upcoming verb?**  
Wing Yee Chow, Cybelle Smith, Ellen Lau, Colin Phillips  
**Animacy modulates cueing strength of mismatching classifiers in marking clausal boundary: ERP evidence from Chinese relatives**  
Fuyun Wu, Yanan Sheng, Xiaolin Zhou  
**Similarity-based interference during comprehension of noun phrases: Evidence from ERPs**  
Andrea E. Martin and Mante S. Nieuwland  
**Integrate or repair? ERP responses to semantic anomalies depend on choice of processing strategy**  
Leif Oines and Albert Kim |
| 15:30  | Coffee                                                                  |
| 16:00  | **Processing reciprocity: What influences the strength of ‘each other’?**  
Elsi Kaiser  
**Language understanding & common sense reasoning**  
Joshua K. Hartshorne, Tim O'Donnell, Joshua Tenenbaum, Tobias Gerstenberg  
**Incremental interpretation and reference to contrast sets**  
Craig Chambers and Daphna Heller  
**Cognitive load affects theory of mind use in the director task**  
Jessica Wang, Steven Frisson, Ian Apperly |
Posters

THURSDAY SEPTEMBER 4

01 Pupillary evidence of ambiguity resolution in Italian subject anaphora
Margreet Vogelzang, Petra Hendriks, Hedderik Van Rijn

02 Bidirectional syntactic priming in conversation: I am primed by you if you are
primed by me
Lotte Schoot, Peter Hagoort, Katrien Segaert

03 The effect of motivation and working memory load on perspective-use:
Evidence from eye-tracking
James Cane, Heather Ferguson, Ian Apperly

04 Category systems gradually align over learning and communication
Catriona Silvey, Simon Kirby, Kenny Smith

05 Investigation of Chinese-English learners’ acquisition of tense and aspect in
English: Self-paced reading and eye-movement study
Panpan Yao and Baoguo Chen

06 Words are less salient in phrases: Evidence from the frequency effect in
recognition memory
Cassandra L. Jacobs, Gary Dell, Aaron Benjamin

07 How does the old support the new? The role of memory consolidation in
learning novel morphological forms
Jelena Mirkovic and Gareth Gaskell

08 Strains and symptoms of the ‘literacy virus’: Modelling the effects of
orthographic transparency on phonological processing
Alastair Smith, Padraic Monaghan, Falk Huettig

09 Misbehaved masculines: Incidental acquisition of gender in L2 German
Denisa Bordag, Amit Kirschenbaum, Andreas Opitz, Maria Rogahn, Erwin Tschirner

10 Phonotactic probability in adult L1 and L2 word learning
Amit Kirschenbaum, Denisa Bordag, Maria Rogahn, Erwin Tschirner

11 Processing of different types of affixes in Czech
Radka Julínková and Denisa Bordag

12 Less isn’t always more: The negative effect of prior exposure in non-adjacent
dependency-learning in adults
Ileana Grama, Annemarie Kerkhoff, Frank Wijnen

13 Resolution of different pronominal forms: French compared to Spanish
Saveria Colonna, Sarah Schimke, Carlos Gelormini-Lezama, Amit Almor

14 Definiteness affects pronoun resolution: Evidence from German
Klaus von Heusinger, Sofiana Chiriacescu, Andreas Brocher

15 Pronoun interpretation is sensitive to verb semantics, except when it isn’t:
Selective verb effects on Italian null and overt pronouns
Emily Fedele and Elsi Kaiser

16 NP status in the establishment of focus and processing of anadeixis
Derya Cokal, Patrick Sturt, Fernanda Ferreira

17 The online comprehension of indirect replies: Evidence from eye-tracking
Jeffrey Wood, Matthew Haigh, Jamal Kinsella, Andrew Stewart
Language prediction requires the same cognitive resource for language comprehension
Aine Ito, Martin Pickering, Martin Corley

Mechanism under auditory perceptual simulation: Speech speed or attitudes biases?
Peiyun Zhou and Kiel Christianson

Resumptive pronouns ameliorate island violations in forced-choice tasks
Lauren Ackerman, Michael Frazier, Masaya Yoshida

Judging paradigmatic relations: A new collection of English ratings
Gabriella Lapesa, Sabine Schulte im Walde, Stefan Evert

Comprehenders pre-activate only the best sentence completion
Daniel Kleinman, Elin Runnqvist, Victor Ferreira

Lexical representations of nouns in German rely on underspecified gender features
Andreas Opitz and Thomas Pechmann

Subliminal morphological speech priming
Roberto Bottini and Davide Crepaldi

Effects of talker and noise variation on recognition memory for spoken words
Angela Cooper and Ann R. Bradlow

Verbs as spatial queues: Explicit direction of motion facilitates visual probe detection
Ashley Chapman and Andriy Myachykov

The (non)effect of discourse relations on comprehending verb phrase ellipsis
Charles Clifton

Diverging online storage costs for wh-adjuncts are predicted by the grammatical theory
Arthur Stepanov and Calum Riach

Sarcasm comprehension: The effects of literality, familiarity, and expectation
Alexandra Turcan and Ruth Filik

Prediction using production or production engaging prediction?
Florian Hintz, Antje Meyer, Falk Huettig

Mandarin subject relatives are easier to process than object relatives even in discourse contexts
Fuyun Wu and Yanan Sheng

Both native and non-native disfluencies trigger listeners’ attention
Hans Rutger Bosker, Jade Tjong, Hugo Quenê, Ted Sanders, Nivja H. de Jong

Tip-of-the-tongue in a second language: The effects of brief L1 exposure on L2 production
Hamutal Kreiner and Tamar Degani

Activations of contextually-appropriate and inappropriate meanings dissociate during semantic reinterpretation of sentences
Sylvia Vitello, Rebecca Tan, Joseph T. Devlin, Jennifer Rodd

Agreement attraction in modern standard Arabic comprehension
Matthew Tucker, Ali Idrissi, Diogo Almeida

Agreement attraction in the neural language system
Matthew Tucker, Stephen Politzer-Ahles, Diogo Almeida

Beyond single word production: How do speakers produce noun phrases?
Zeshu Shao

Development in perspective-taking during ambiguity avoidance
Kumiko Fukumura, Lauren Ferguson, Katie Quinn
Syntactic priming: A lexical boost, cumulativity, an inverse preference effect and... A positive preference effect
Katrien Segaert and Peter Hagoort

Effects of neighbourhood density on lexicality judgements across the lifespan and in individuals with Alzheimer’s disease
Nancy Azevedo, Eva Kehayia, Ruth Ann Atchley, Vasavan N.P. Nair

How language mediates the creation of false memories
Agneszka Konopka

Fast and furious lexical processing
Jacolien van Rij, Cyrus Shaoul, R. Harald Baayen, Petar Milin

Anomalous transfer of syntactic rules between languages
Awel Vaughan-Evans, Jan Kuipers, Guillaume Thierry, Manon Jones

A neurocomputational model of the N400 and the P600
Harm Brouwer and John Hoeks

Word frequency in context shows differential effects on eye fixations and fixation-related potentials
Franziska Kretzschmar, Matthias Schlesewsky, Adrian Staub

Recovering from syntactic and semantic violations: The relationship between eye movements and brain responses
Paul Metzner, Titus von der Malsburg, Shravan Vasishth, Frank Rössler

Marking the counterfactual: ERP evidence for pragmatic processing of German subjunctives
Eugenia Kulakova, Dominik Freunberger, Dietmar Roehm

On the impact of phrasal length on native and non-native language production: Evidence from Basque and Spanish
Idoia Ros, Mikel Santesteban, Kumiko Fukumura, Itziar Laka

Amplitude modulated noise negatively affects processing of prosodic cues
Rebecca Carroll and Esther Ruigendijk

The role of orthography in recognizing connected speech
Holger Mitterer and Eva Reinisch

What was the question? Broad focus in focus-sensitive coordination
Jesse Harris and Katy Carlson

Phrase lengths and prosody in pp-attachment
Ronit Webman Shafran and Janet Dean Fodor

Accents influence attachment generally
Katy Carlson

The resolution of the clause that is relative: Prosody and plausibility in RC attachment
Daniela Zahn and Christoph Scheepers

A sketch of an extrinsic timing model of speech production
Alice Turk and Stefanie Shattuck-Hufnagel

Prosody determines syntactic choice in language production and reading
Gerrit Kentner

Stress assignment in the development of Italian reading ability: Effects of lexical stress, frequency, and stress neighbourhood
Simone Sulpizio and Lucia Colombo

When exactly corners stop corning? Incremental masked priming and complex word identification
Davide Crepaldi, Marco Marelli, Simona Amenta, Anais De Santis, Elena Angela Morone, Rocco De Marco
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
<td><strong>Distributive and collective readings of plural DPs</strong></td>
<td>Carmen Dobrovie-Sorin, Emilia Ellsiepen, Barbara Hemforth</td>
</tr>
<tr>
<td>60</td>
<td><strong>Between-group matching of confounding variables: Why covariates remain important for analysis</strong></td>
<td>Christoph Scheepers</td>
</tr>
<tr>
<td>61</td>
<td><strong>Co-indexing beyond nouns and verbs: The link between visuo-spatial information and abstract adverbs of manner</strong></td>
<td>Ernesto Guerra and Pia Knoeferle</td>
</tr>
<tr>
<td>62</td>
<td><strong>The properties of people's social circle affect their semantic skills</strong></td>
<td>Shiri Lev-Ari</td>
</tr>
<tr>
<td>63</td>
<td><strong>Training and timing local enrichments</strong></td>
<td>Emmanuel Chemla, Chris Cummins, Raj Singh</td>
</tr>
<tr>
<td>64</td>
<td><strong>The role of numerical cognition in signalling question under discussion</strong></td>
<td>Chris Cummins</td>
</tr>
<tr>
<td>65</td>
<td><strong>Alternative sets in language processing: The mechanisms of activation and competitive inhibition</strong></td>
<td>Nicole Gotzner, Isabell Wartenburger, Katharina Spalek</td>
</tr>
<tr>
<td>66</td>
<td><strong>Embodied conceptualization in language production</strong></td>
<td>Zhenguang Cai</td>
</tr>
<tr>
<td>67</td>
<td><strong>Instrument priming: A tale of two strategies</strong></td>
<td>Chung-I Erica Su and Gail Mauner</td>
</tr>
<tr>
<td>68</td>
<td><strong>Categorization and implicit learning of semantic roles</strong></td>
<td>Eva Wittenberg, Manizeh Khan, Jesse Snedeker</td>
</tr>
<tr>
<td>69</td>
<td><strong>Processing modifier attachment ambiguities: Evidence from Russian</strong></td>
<td>Daria Chernova and Natalia Slioussar</td>
</tr>
<tr>
<td>70</td>
<td><strong>The effect of gradient foreign accentedness and listener experience on lexical activation strength and access</strong></td>
<td>Vincent Porretta, Juhani Järvikivi, Benjamin V. Tucker</td>
</tr>
<tr>
<td>71</td>
<td><strong>Listeners recognize others’ speech better than their own</strong></td>
<td>William Schuerman, Antje Meyer, James McQueen</td>
</tr>
<tr>
<td>72</td>
<td><strong>Phonological representations in adults with dyslexia: Evidence from a perceptual learning paradigm</strong></td>
<td>Margriet Groen and James McQueen</td>
</tr>
</tbody>
</table>
01 **Overspecification in reference: Modelling size contrast effects**  
Roger van Gompel, Albert Gatt, Emiel Krahmer, Kees van Deemter

02 **An information-theoretic account of word memorability**  
Kyle Mahowald, Phillip Isola, Evelina Fedorenko, Aude Oliva, Edward Gibson

03 **Common ground and joint utterance production: Evidence from the word chain task**  
Jaroslav Lelonkiewicz and Chiara Gambi

04 **Conceptual causes and consequences of label choice in dialogue**  
Eugene Philalithis, Anita Böröndi, Holly Branigan

05 **The relationship between implicit expectations about character behaviour and eating disorder tendencies: Evidence from eye movements during reading**  
Christina Ralph-Nearman and Ruth Filik

06 **The earlier, the better? Bilingual cognitive control and age of acquisition**  
Santa Vinerte and Laura Sabourin

07 **Susceptibility of language mechanisms to input over development**  
Marius Janciauskas and Franklin Chang

08 **Young children represent syntax abstractly, but prefer to use it concretely**  
Nick Gruberg, Liane Wardlow, Victor Ferreira

09 **Multiple syntactic processing biases that change over development**  
Kirsten Abbot-Smith, Franklin Chang, Caroline Rowland, Heather Ferguson, Julian Pine

10 **The role of different types of cues in disambiguating non-canonical sentences**  
Valentina Cristante

11 **The processing of morphosyntactic cues in wh-questions: An eye-tracking study**  
Atty Schouwenaars, Esther Ruigendijk, Petra Hendriks

12 **The acquisition of L1-L2 grammatical gender in German-French and English-French bilinguals**  
Leah Pappas, Carrie Jackson, Janet Van Hell, Katharina Spalek

13 **Cognitive functioning in late unbalanced bilingualism and multilingualism**  
Mariana Vega-Mendoza, Antonella Sorace, Thomas H. Bak

14 **Effects of animacy and semantic relatedness in native and non-native sentence processing: An ERP study**  
Mariana Vega-Mendoza, Martin Pickering, Mante Nieuwland

15 **Correlations in hemodynamic fluctuations between cortical regions during non-native phonemic category processing among participants from late childhood to young adulthood**  
Kiyomi Yatabe, Ei-ichi Hoshino, Reiko Akahane-Yamada, Nozomi Naoi, Yasuyo Minagawa

16 **Pitch accent type and metrical stress in infant speech segmentation**  
Bettina Braun, Muna Pohl, Katharina Zahner

17 **Interactive participation facilitates children’s processing of direct speech reports**  
Franziska Köder

18 **Orthographic processing in bilingual children: The role of semantics in cognate recognition**  
Pauline Schröter
19 The impact of task context on children’s acquisition of inflection: Experimental evidence and computational simulation
Moy Joanna, Anna Theakston, Grzegorz Krajewski, Anna Woollams

20 Perspective-taking in text is modulated by how information is presented (direct vs. indirect speech)
Steven Frisson, Elisa Back, Alice Rees, Ian Apperly

21 Irony and emoticons: Comprehension and emotional impact
Ruth Filik, Alexandra Turcan, Dominic Thompson, Nicole Harvey, Harriet Davies, Hartmut Leuthold

22 Which is stronger - gaze versus the recent event preference?
Dato Abashidze, Pia Knoeferle, Maria Nella Carminati

23 Knowledge of what others know constrains discourse processing
Yuki Kamide and Anuenue Kukona

24 Spatial influences on object accessibility in discourse processing.
Glenn Williams, Anuenue Kukona, Yuki Kamide

25 Does pronoun processing vary across languages?
Sol Lago, Shayne Sloggett, Zoe Schlueter, Wing Yee Chow, Ellen Lau, Colin Phillips

26 Comprehenders infer interaction between meaning intent and grammatical probability
Mark Myslin and Roger Levy

27 The processing of morphologically complex words in Turkish heritage speakers
Gunnar Jacob and Bilal Kırkıcı

28 Effects of local predictability on eye fixation behavior in silent and oral reading for younger and older adults
Cornelia Moers, Antje Meyer, Esther Janse

29 The effect of exposure in cross-dialect perception: Hearing ambiguous /r/ variants in Glaswegian
Robert Lennon

30 Directionality and underspecification: The processing of zero-derivation
Jeannique Darby

31 Inhibitory interference in reflexives: Evidence for cue confusability
Lena Jäger, Felix Engelmann, Shravan Vasishth

32 Early interference effects of non-structural gender: Behavioral evidence from Chinese pronouns and complex reflexives
Jingjing Qian and Fuyun Wu

33 Plausibility and filled gap effects in native and non-native sentence processing
Ian Cunnings and Claudia Felser

34 WH-dependency resolution interrupts reflexive antecedent search on-line
Michael Frazier, Peter Baumann, Lauren Ackerman, David Potter, Masaya Yoshida

35 Agreement inhibits detection of spray/load mismatches in gapping
Michael Frazier and Masaya Yoshida

36 Saturating the anticipatory mechanisms of syntactic priming during situated language understanding
Eunice Fernandes, Moreno I. Coco, Holly P. Branigan

37 Working memory, inhibition, and grammatical role assignment
Alix Kowalski and Yi Ting Huang

38 Readers with less cognitive control are more affected by surprising content
Bruno Nicenboim, Shravan Vasishth, Reinhold Kliegl
39 Certain people don’t need modification: The effect of indefiniteness on Japanese relative clause association preferences
Chunhua Bai, Doug Roland, Yuki Hirose

40 Shared processing of linguistic and musical structure: Evidence from interference effects
Joris Van de Cavey and Robert Hartsuiker

41 Children’s online processing of subject and object relative clauses in Greek
Kalliopi Katsika and Shanley Allen

42 Reactivation effects interact with expectation strength
Samar Husain and Shravan Vasishth

43 Priming prosody: Speech rate and boundary placement
Kristen Tooley, Agnieszka Konopka, Duane Watson

44 Prediction gone wrong: A mouse-tracking study in sentence comprehension
Franklin Chang and Agnieszka Konopka

45 Pre-attentive processing of consonantal duration contrasts
Sandra Kotzor, Adam C. Roberts, Allison Wetterlin, Aditi Lahiri

46 Individual differences in word recognition in noise
Jens Schmidtke

47 Integration of “secondary” information during comprehension: Co-speech gestures and mouth patterns accompanying signs
David Vinson, Pamela Perniss, Wing-San Chan, Amy Rice, Gabriella Vigliocco

48 Cognitive cost of switching between standard and dialect varieties
Neil Kirk, Mathieu Declerck, Ken Scott-Brown, Vera Kempe, Andrea Philipp

49 Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences
Inbal Arnon and Uriel Cohen Priva

50 (Mis)binding of thematic roles and grammatical functions in passive sentences
Heeju Hwang

51 Is explicit memory for prime sentence structure necessary for lexically based syntactic priming?
Sarah Bernolet, Paul Boon, Miet De Letter, Robert Hartsuiker

52 Are thematic roles linked to grammatical function and linear order in the same way? Evidence from structural priming effects in Japanese sentence production
Ying Deng, Douglas Roland, Yuki Hirose

53 An electrophysiological analysis of the time course of phonological and orthographic encoding in written word production
Qingqing Qu and Markus Damian

54 Language-wide inhibition? Maybe not! Unbalanced bilinguals exhibit balanced switch costs for high-frequency targets
Daniel Kleinman and Tamar Gollan

55 The influence of script knowledge on expectations: Evidence from ERPs
Francesca Delogu, Heiner Drenhaus, Matthew Crocker

56 Retrieval interference during comprehension of grammatical subject-verb agreement: Evidence from ERPs
Mante S. Nieuwland, Ian Cunnings, Andrea E. Martin, Patrick Sturt

57 “What is concrete and abstract and read all over!?” An fMRI study of type coercion in inherent polysemy
Yuan Tao, Andrew Anderson, Massimo Poesio
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>Hemispheric sensitivity to frequency and predictability: Evidence from early and late ERPs</td>
<td>Yoana Vergilova, Heiner Drenhaus, Matthew Crocker</td>
</tr>
<tr>
<td>59</td>
<td>Out of use, out of mind? First language processing in a Dutch emigrant population in Anglophone North America</td>
<td>Bregtje Seton, Susanne Brouwer, Laurie Stowe, Monika Schmid</td>
</tr>
<tr>
<td>60</td>
<td>Proficient bilinguals' lexical access is affected by the context of interlocutor identity</td>
<td>Monika Molnar, Clara Martin, Antonio Ibañez, Manuel Carreiras</td>
</tr>
<tr>
<td>61</td>
<td>Transitivity alternations in English: Processing subcategorization frames</td>
<td>Gillian Ramchand, Bjorn Lundquist, Martin Corley, Mai Tungseth, Antonella Sorace</td>
</tr>
<tr>
<td>62</td>
<td>Spatial attention and implied speed in sentence comprehension: Path and adverb manipulations</td>
<td>Shane Lindsay, Christoph Scheepers, Yuki Kamide</td>
</tr>
<tr>
<td>63</td>
<td>Orthography-semantics consistency: A new measure and its impact on visual word recognition</td>
<td>Marco Marelli, Simona Amenta, Davide Crepaldi</td>
</tr>
<tr>
<td>64</td>
<td>The influence of lexical association on syntactic analysis: Eye movement evidence</td>
<td>Anna Fiona Weiß, Franziska Kretzschmar, Ina Bornkessel-Schlesewsky, Adrian Staub</td>
</tr>
<tr>
<td>65</td>
<td>Does a prediction benefit outlast grammatical violations? A double-violation ERP study.</td>
<td>Dominik Freunberger</td>
</tr>
<tr>
<td>66</td>
<td>The misinterpretation of noncanonical sentences revisited</td>
<td>Markus Bader and Michael Meng</td>
</tr>
<tr>
<td>67</td>
<td>Semantic interference effects in picture naming: The effect of manipulating type of semantic relations on name retrieval times</td>
<td>Melanie Vitkovitch, Anita Potton, Hisma Begum, Aimen Hasham, Mary-Maria Johnson</td>
</tr>
<tr>
<td>68</td>
<td>Limits to cross-modal semantic and object shape priming in sentence context</td>
<td>Joost Rommers and Falk Huettig</td>
</tr>
<tr>
<td>69</td>
<td>Before and after, and processing presuppositions in discourse</td>
<td>Ming Xiang, Emily Hanink, Genna Vegh</td>
</tr>
<tr>
<td>70</td>
<td>Prediction of optional heads for maintaining incrementality of parsing</td>
<td>Nina Kazanina</td>
</tr>
<tr>
<td>71</td>
<td>Cascaded lexical processing to semantics in spoken embedded words</td>
<td>Jenna Todd Jones, Jeffrey Bowers, Nina Kazanina</td>
</tr>
</tbody>
</table>
POSTERS

SATURDAY SEPTEMBER 6

01 From guessing to getting it right: The initial development of an L2 lexicon
Geertje van Bergen, Béryl Hilberink-Schulpen, Ton Dijkstra

02 Putting things in new places: Verb-based prediction in L1 and L2 sentence processing
Monique Flecken and Geertje van Bergen

03 “Right now, Sophie *swims in the pool?!” Processing of grammatical aspect in native and second language readers
Kelly Walbert and Monique Flecken

04 When growing a fruitless tree is a fruitless effort: Context modulation of the processing of morpheme meanings in complex words
Simona Amenta, Marco Marelli, Davide Crepaldi

05 Structural alignment is greater in dialogue than in monologue
Iva Ivanova and Victor Ferreira

06 Social influences on the regularization of unpredictable linguistic variation
Olga Feher, Simon Kirby, Kenny Smith

07 The power of imageability: Effects of semantic factors on the acquisition of inflections
Filip Smolik and Adam Kríž

08 The development of eye-movements in three different orthographies: An eye-tracking investigation
Sascha Schroeder, Tuomo Häikiö, Ascen Pagan, Jonny Dickins, Jukka Hynölä

09 Limits and variations of linguistic generalizations
Silvia Radulescu, Frank Wijnen, Sergey Avrutin

10 The bouba effect: Sound-shape iconicity is based on roundness
John Matthew Jones, David Vinson, Nourane Clotsté, Alex Lau Zhu, Julio Santiago, Gabriella Vigliocco

11 Sensitivity to prosodic cues and thematic emphasis in children’s sentence continuation
Mirta Vernice and Maria Teresa Guasti

12 Subject-verb agreement attraction in production by Japanese learners of English
Mayuko Yusa, Jungho Kim, Noriaki Yusa, Masatoshi Koizumi

13 How persistent is syntactic priming? Exploring long-term priming effects in children and adults
Katherine Messenger

14 Computational modeling of acquisition of vocal imitation using interaction between a virtual infant and a human participant
Heikki Rasilo

15 Tracking emerging sensitivity to L2 morphosyntax: Evidence for the role of the L1 from two longitudinal studies of early learners
Alison Gabriele, Robert Fiorentino, Jose Aleman Banon, Lauren Covey, Amy Rossonomo

16 Memory retrieval favors referential antecedents: Interference effects in the processing of reflexives
Peter Baumann and Masaya Yoshida

17 Number attraction effects on object-clitic agreement in Spanish: Evidence from native and non-native speakers
Adam Zawiszewski and Mikel Santesteban
POSTERS

SATURDAY SEPTEMBER 6

18 A comparison of music and language reading: Rapid eye-movement effects of musical incongruence
Lauren Hadley, Patrick Sturt, Martin Pickering

19 Thematic fit does not always help ambiguity resolution
Manabu Arai, Chie Nakamura, Yuki Hirose

20 Getting ready for a surprise
Manabu Arai, Yuki Hirose, Chie Nakamura

21 Modeling visual world eye-tracking data: Insights from fMRI
Emily Atkinson, Akira Omaki, Colin Wilson

22 Grammatical faithfulness in processing of multiple dependencies
Emily Atkinson and Akira Omaki

23 Case and finiteness versus structural position in subject-verb attachment
Nathan Arnett and Matt Wagers

24 Eye movements during the processing of pronominal relative clauses
Douglas Roland, Yuki Hirose, Gail Mauner, Stephani Foraker

25 The expectation effect of roles and role fillers in Korean: Evidence from ERP and self-paced reading studies
Hongoak Yun, Dongsu Lee, Hey-in Jeong, Upyong Hong

26 Temporal dynamics of extraction from WH-islands: A speed-accuracy trade-off study
Sandra Villata, Brian Mcelree, Julie Franck

27 Navigating interference effects in agreement in sentence comprehension
Julie Franck, Maria Lucia Cali, Luigi Rizzi

28 Small clause availability triggers high attachment in English
Nino Grillo, Bruno Fernandes, João Costa, Andrea Santi

29 Comprehension priming evidence for elliptical structures
Julian Grove, Emily Hanink, Ming Xiang

30 Adding a third WH-phrase sometimes increases the acceptability of multiple-WH-questions
Lauren Ackerman and Masaya Yoshida

31 Repeated exposure to a word meaning protects against interference from exposure to its alternative meaning
Hannah Betts, Zhenguang Cai, Jennifer Rodd

32 Disfluency and attention in language comprehension
Sam Miller, Martin Corley, Martin Pickering, Ian Finlayson, Eleanor Drake

33 Elaboration of targets and non-targets in online sentence comprehension
Philip Hofmeister and Shravan Vasishth

34 Participants exploit dependencies hidden in studies employing latin square designs
Jakub Szewczyk

35 English resumptive pronouns are predicted by bare-bones syntax
Adam Morgan

36 Working memory and distance in language production
Markus Bader

37 Priming from music to language: Effects of number marking on the priming of relative clause attachment in Spanish
Mythili Menon and Elsi Kaiser

38 Parafoveal preview benefit during second language reading
Chi Yui Leung and Masatoshi Sugiura
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>No morphology in the parafovea: Evidence from an invisible-boundary</td>
<td>Federica Degno and Davide Crepaldi</td>
</tr>
<tr>
<td></td>
<td>eye-tracking experiment</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Independent effects of beat gesture and pitch accent on processing</td>
<td>Diana Dimitrova, Mingyan Chu, Lin Wang, Asli Özyürek, Peter Hagoort</td>
</tr>
<tr>
<td></td>
<td>words in context</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>On-line processing of moral transgressions: ERP evidence for</td>
<td>Hartmut Leuthold, Angelika Kunkel, Ian Grant Mackenzie, Ruth Filik</td>
</tr>
<tr>
<td></td>
<td>spontaneous evaluation</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Examining the neural correlates of irony and sarcasm: An fMRI study</td>
<td>Ruth Filik, Alexandra Turcan, Christina Ralph-Nearman, Alain Pitiot</td>
</tr>
<tr>
<td>43</td>
<td>Modeling neural correlates of syntactic structure building</td>
<td>John Hale, Frederick Callaway, Elana Feldman, Jaclyn Jeffrey-Wilensky, David Lutz, Adam</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mahar, Jonathan Brennan</td>
</tr>
<tr>
<td>44</td>
<td>Recasting the eLAN as an attentional efficiency-dependent enhancement</td>
<td>Christopher Barkley, Robert Kluender, Marta Kutas</td>
</tr>
<tr>
<td></td>
<td>of the N100</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>From case markers to phrase structures: ERP evidence on predictive</td>
<td>Yunju Nam, Hongoak Yun, Upyong Hong, Eunyong Chung</td>
</tr>
<tr>
<td></td>
<td>sentence processing in Korean</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Working memory influences processing of whether-islands: The N400</td>
<td>Dan Michel, Robert Kluender, Seana Coulson</td>
</tr>
<tr>
<td></td>
<td>reflects gap predictability in high span readers</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>The downstream consequences of form-to-expectation mismatch</td>
<td>Emily S. Nichols, Marc F. Joanisse, Thomas A. Farmer</td>
</tr>
<tr>
<td>48</td>
<td>CLIQS: Cross-linguistic investigations in quantitative syntax</td>
<td>Richard Futrell, Kyle Mahowald, Edward Gibson</td>
</tr>
<tr>
<td>49</td>
<td>The role of global syntactic context in structural priming: Evidence</td>
<td>Gunnar Jacob, Kalliopi Katsika, Alina Kholodova, Neiloufar Family, Shanley Allen</td>
</tr>
<tr>
<td></td>
<td>from German</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Resolving lexical ambiguity using sub-phonemic duration cues: Did you</td>
<td>Yoonjeong Lee, Elsi Kaiser, Louis Goldstein</td>
</tr>
<tr>
<td></td>
<td>hear “place kin” or “play skin”?</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>The recognition of assimilated forms during real-time comprehension:</td>
<td>Mercedes Mohaghegh and Craig Chambers</td>
</tr>
<tr>
<td></td>
<td>Effects of phonological context and priming</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Production of noun phrases containing allophonic word-onsets</td>
<td>Emily Darley, Markus Damian, Nina Kazanina</td>
</tr>
<tr>
<td>53</td>
<td>Phonological units of speech production in Mandarin-English</td>
<td>Laura Burchfield and Ann Bradlow</td>
</tr>
<tr>
<td></td>
<td>bilinguals: Variable or fixed strategies?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>Parafoveal processing of phonological and orthographic information</td>
<td>Simon Tiffin-Richards and Sascha Schroeder</td>
</tr>
<tr>
<td></td>
<td>in children and adults during sentence reading</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Prosodic prominence: Information status or production fluency?</td>
<td>Jennifer Arnold and Elise Rosa</td>
</tr>
<tr>
<td>56</td>
<td>The time course of semantic constraints in pronoun resolution</td>
<td>Israel de la Fuente and Barbara Hemforth</td>
</tr>
<tr>
<td>57</td>
<td>Information density and the use of discourse cues</td>
<td>Fatemeh Torabi Asr and Vera Demberg</td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Authors</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>58</td>
<td><strong>Much of flies and insects. Cross-language activation and inhibition in interlingual homographs comprehension.</strong></td>
<td>Joanna Durlik, Jakub Szewczyk, Zofia Wodniecka</td>
</tr>
<tr>
<td>59</td>
<td><strong>No squirrel likes collecting nuts. An ERP study on quantification, prediction, and cloze probability.</strong></td>
<td>Dominik Freunberger and Dietmar Roehm</td>
</tr>
<tr>
<td>60</td>
<td><strong>Incremental processing of quantificational restriction: Evidence from cross-modal probe recognition</strong></td>
<td>Petra Augurzky, Oliver Bott, Wolfgang Sternefeld, Rolf Ulrich</td>
</tr>
<tr>
<td>61</td>
<td><strong>Looking beyond accessible referents. Effects of different types of indefinite noun phrases on the subsequent discourse</strong></td>
<td>Sofiana Chiriacescu</td>
</tr>
<tr>
<td>62</td>
<td><strong>Hemispheric differences in processing the scenario relevance of words in discourse</strong></td>
<td>Ross Metusalem, Marta Kutas, Thomas P. Urbach, Jeffrey L. Elman</td>
</tr>
<tr>
<td>63</td>
<td><strong>The truth before and after: Temporal connectives modulate online sensitivity to truth-value</strong></td>
<td>Mante S. Nieuwland</td>
</tr>
<tr>
<td>64</td>
<td><strong>How vertical hand movements impact brain activity elicited by literally and metaphorically related words</strong></td>
<td>Seana Coulson and Megan Bardolph</td>
</tr>
<tr>
<td>65</td>
<td><strong>Effects of working memory load on the processing of numerals and standard scalar quantifiers</strong></td>
<td>Laine Stranahan, Qingqing Wu, Jesse Snedeker, Manizeh Khan</td>
</tr>
<tr>
<td>66</td>
<td><strong>Gaps within fillers are not better than gaps within subjects</strong></td>
<td>Grant Goodall</td>
</tr>
<tr>
<td>67</td>
<td><strong>Psycholinguistic evidence for “plural” as a value of gender in Konso</strong></td>
<td>Mulugeta Tsegaye, Maaren Mous, Niels Schiller</td>
</tr>
</tbody>
</table>
ELIMINATING UNPREDICTABLE LINGUISTIC VARIATION THROUGH INTERACTION
Kenny Smith (University of Edinburgh), Olga Fehér (University of Edinburgh) & Nikolaus Ritt (University of Vienna)
kenny.smith@ed.ac.uk

Natural languages do not differ arbitrarily, but are constrained so that certain properties recur across languages. These universals presumably arise from the interaction between the processes of language learning and language use (see e.g. Christiansen & Chater, 2008). A test-case for the relationship between these factors is linguistic variation (e.g. Hudson Kam & Newport, 2005). Variation in language tends to be predictable: usage of alternate forms is conditioned in accordance with phonological, semantic, pragmatic or sociolinguistic criteria. Previous experimental work shows that language learning (either by single individuals, as in Hudson Kam & Newport, 2005, or across multiple episodes of iterated learning, as in Reali & Griffiths, 2009, Smith & Wonnacott, 2010) leads to the loss of unpredictable variation, showing that language transmission from individual to individual generates the kind of predictable variability seen in language. Here we consider a second mechanism by which variation might be systematically eliminated or conditioned: alignment during interaction. Interlocutors modify their linguistic behaviour towards that of their partners (a process known as accommodation or alignment, see e.g. Pickering & Garrod, 2004; Coupland, 2010). These low-level adjustments potentially impact on the structure of language, both by changing the long-term behaviour of the individuals involved and by skewing the linguistic data from which other individuals learn.

We ran an experiment, completed in pairs, in which participants were asked to learn a miniature language, attempted to reproduce that language during an initial test phase, and then used it to communicate with their partner during a second test phase. The miniature language was used to describe scenes involving animals performing movements: scenes differed in the animal(s) involved, the movement involved, and the number of animals in the scene (one or two). We manipulated the proportion of training trials on which singular number was redundantly marked with an overt linguistic marker, a post-nominal particle (plural number was always overtly marked). We were interested in how participants changed their use of this singular marker during interaction.

We label conditions according to the language pairs were trained on: e.g. in the 66-33 condition, for one participant singulars were marked in 66% of their training data, for the other participant singulars were marked 33% of the time. In two conditions where both members of a pair were trained on a language which marked the singular variably (conditions 66-33 and 87-17), participants accurately reproduced the variability of their input language during initial testing, but rapidly aligned during interaction with their partner on a language which exhibited no variation or conditioned variation (i.e. always or never marking singular overtly, or only for certain nouns). We ran two further conditions (100-66 and 100-33) in which one member of each pair was trained on a non-varying, categorical system. Categorically-trained and variably-trained participants differed in the extent to which they accommodated to their partner during interaction: variably-trained participants changed their marker usage significantly more than categorical participants, and reliably increased their usage towards that of their partners. Comparison with 66-33 and 87-17 conditions shows that these results are due to interacting with a categorical user.

Alignment during interaction leads to the elimination of unpredictable variation, and provides a mechanism by which the absence of unpredictable variation in natural language might be explained. Furthermore, accommodation during interaction is inherently asymmetric: while variable users can accommodate to a categorical partner by increasing their frequency of usage, categorical users tend not to accommodate to their variable partners by becoming variable. This finding on alignment during language use has implications for our understanding of language change: once a grammatical marker reaches a critical threshold in a population such that many individuals are variable and some are categorical, alignment during interaction should drive the population towards uniform categorical marker use.
INTERFERENCE DURING ANTECEDENT RETRIEVAL: EFFECTS OF READING SKILL
Dave Kush (Haskins Laboratories)
kush@haskins.yale.edu

A growing body of research has investigated the real-time application of binding constraints by asking whether feature-matching distractors in syntactically inappropriate positions interfere with the computation of anaphoric dependencies (e.g. Badecker & Straub 2002; Sturt 2003). Although many studies have found that comprehenders often successfully apply binding constraints in early processing, recent findings suggest that susceptibility to interference may vary across populations (e.g. Cunnings & Felser, 2013). In order to contribute to a general understanding of variability in successful application of binding constraints online we present results from two eye-tracking studies that investigate interference among skilled readers (college undergraduates, Experiment 1, N=33) and less-skilled readers (age-matched non-college-going individuals, Experiment 2, N=60).

Our binding constraint of interest in the experiments is the Scope Constraint on bound-variable (BV) pronouns, which prohibits a Quantificational NP (QP) from binding any pronoun not in its scope (e.g. Reinhart 1983; Barker 2012). Our experiments manipulated two factors: (i) the GENDER-MATCH between an embedded RC-subject (girl/boy scout) and the critical pronoun (her) and (ii) the RC-subject’s REFERENTIALITY. When the RC-subject was a QP, as in (1), it was not an eligible antecedent for the pronoun because of the Scope Constraint, regardless of gender-match. When the RC-subject was referential (as in 2), gender-match determined its ability to serve as a co-referential antecedent for the pronoun.

(1) The troop leaders [that no girl/boy scout had any respect for] had scolded her after the incident at the scout camp.

(2) The troop leaders [that the girl/boy scout had any respect for] had scolded her after the incident at the scout camp.

We expected that gender-match should facilitate processing in (2), where a grammatical relationship could be established between RC-subject and pronoun. A comparable effect of gender-match in the Quantificational conditions would indicate interference (i.e. failure to ignore the inaccessible QP). If, however, comprehenders can successfully ignore a feature-matching distractor, gender-match should have no effect in Quantificational conditions.

In Experiment 1, gender-match facilitated processing at the pronoun only when the RC-subject was referential (Rightbound RTs, t = -2.40). No facilitation was observed when the RC-subject was quantificational. This GenderMatch x Referentiality interaction suggest that college undergraduates display immediate sensitivity to the BV constraint and a lack of interference. Non-college-going individuals showed a contrasting pattern of effects (Experiment 2). Processing of the region directly following the pronoun was facilitated when the RC-subject matched the pronoun in gender features (First Pass RTs: t = 2.12; RBRT: t = 4.15). The Gender-match x Referentiality interaction observed in Experiment 1 emerged in the post-pronoun region in Regression-Path Duration (t = -2.26).

In sum, both groups showed effects consistent with sensitivity to the grammatical constraint on BV licensing. However, the groups differed in (i) the time-course of the effect, and (ii) whether a period of interference preceded grammatical sensitivity. College-going participants demonstrated immediate sensitivity to the BV constraint, whereas the non-college-going participants demonstrated initial susceptibility to interference from distractors and later alignment with the constraint. Taken together, the results suggest that susceptibility to interference may vary as a function of reading skill. We hypothesize that these effects may reflect a difference in choice of antecedent retrieval cues between the populations. Less skilled readers appear to prioritize gender information, whereas skilled readers favor structural cues that are less interference-prone (e.g. Dillon et al. 2013).

ARE THERE UNIVERSAL PROSODIC CUES TO SPEECH STRUCTURE?
Laurence White (Plymouth University), Sven Mattys (University of York), Silvia Benavides-Varela (IRCCS San Camillo Hospital, Venice; LPP CNRS-Université Paris Descartes) and Katalin Mády (Hungarian Academy of Sciences, Budapest)
laurence.white@plymouth.ac.uk

Prosodic variation informs listeners about speech structure. Furthermore, perceptual interpretation of the valency of variation along prosodic dimensions appears universal: sounds that are longer, louder or higher in pitch are more salient, other things being equal. Of course, languages differ in how prosodic features are combined and organised with respect to lexical, syntactic and pragmatic structure: for example, patterns of stress placement and the interpretation of intonational sequences are well documented to be language-specific (e.g., Ladd, 2008). However, there is evidence that certain form-function associations are universal. For example, the iambic-trochaic law asserts that elements made salient through higher pitch or greater loudness are interpreted as sequence-initial, whilst lengthened segments are typically interpreted as sequence-final (Hayes, 1995).

The use of final vowel lengthening as a cue to speech structure has been demonstrated for several languages and may well be universal (e.g., Beckman, 1992; Tyler & Cutler, 2009). Lengthening of onset consonants in word-initial syllables is widely observed across languages, however (Keating, Cho, Fougeron & Hsu, 2003), and is also a potentially universal segmentation cue. We examined whether the localisation of lengthening – consonant vs vowel – is indeed critical for its structural interpretation, and whether this interpretation is language-independent.

Using an artificial language learning paradigm (Saffran, Newport & Aslin, 1996), we tested how durational variation in sequences of nonsense trisyllables affected listeners' segmentation behaviour. In Experiment 1, synthesised words such as pabiku, golatu, tinudo and daropi were presented in continuous randomised sequence for six minutes, in a between-subjects design with four timing conditions. When the onset consonant of the word-initial syllable was lengthened (e.g., pabiku), recognition in subsequent two-alternative forced-choice tests (words vs foils) was better than both the evenly-timed condition ($p < .05$) and when a medial consonant was lengthened (e.g., pabiku, $p < .005$). Furthermore, lengthening of the initial syllable’s vowel (e.g., pabiku) led to recognition no better than the evenly-timed condition, and worse than with initial consonant lengthening ($p < .01$). This indicates that segmentation is not simply boosted because rhythmic alternation confers a regular timing contrast. Rather, the localisation of the lengthening effect is crucial. Thus, whilst vowel lengthening is a cue to an upcoming boundary (e.g., Saffran et al., 1996), consonant lengthening signals a preceding boundary, at least for English listeners.

Experiment 2, with two further conditions, supported this structural interpretation. When the initial consonant and final vowel were lengthened (e.g., pabiku), recognition was better than for even timing ($p < .001$), while recognition was worse with lengthening of medial vowel-consonant sequences (e.g., pabiku) than in all other conditions ($p < .05$ for all comparisons).

Experiment 3 examined whether the interpretation of localised consonant lengthening as sequence-initial generalised to languages – Hungarian and Italian – with prosodic characteristics rather distinct from those of English. We created a new artificial language in which trisyllabic sequences (bitusa, nudopa, ripolu, subako) did not resemble any real English, Hungarian or Italian words. We found that native speakers of all three languages showed better recognition when initial consonants were lengthened relative to the evenly-timed condition ($p < .001$). Furthermore, there was no interaction between timing condition and native language, suggesting that interpretation of lengthened consonants as sequence-initial may indeed be language-independent.
Neuropsychological investigations of brain-damaged individuals have long suggested that the right hemisphere (RH) plays an important role in processing non-literal aspects of language (e.g., Critchley, 1962; Eisenson, 1962; Zaidel, 1985; Myers, 1999; Lehman Blake, 2005). For example, patients with RH damage experience difficulties with processing conversational inferences (e.g., Kaplan et al., 1990) – including indirect requests (e.g., Weylman et al., 1989; Stemmer, 1994) and commands (e.g., Foldi, 1987) – humor and sarcasm (e.g., Gardner, 1975), and information conveyed by prosody (e.g., Heilman et al., 1975). However, the RH contains several distinct systems that could be contributing to these deficits, including i) the RH homologues of the high-level language processing regions (e.g., Binder et al., 1997; Fedorenko et al., 2010); ii) the RH subset of the system that supports social cognition / Theory of Mind (e.g., Saxe & Kanwisher, 2003); and iii) the RH subset of the domain-general fronto-parietal cognitive control system implicated broadly in goal-directed behavior (e.g., Duncan, 2010). It is therefore difficult to determine which of these RH systems is the key contributor to our pragmatic abilities. This is especially true given that many pragmatic phenomena are complex and thus possibly require some combination of linguistic, social, and generic problem-solving abilities.

We here address this question using the functional localization approach in fMRI. In each participant (n=12), we functionally identified three sets of brain regions using independent “localizer” tasks: i) the language regions (Fedorenko et al., 2010), ii) the regions that support Theory of Mind (ToM; Saxe & Kanwisher, 2003), and iii) the cognitive control regions (Fedorenko et al., 2013). We then examined the responses of these sets of regions in the RH to jokes and their literal controls matched for various lexical-level factors known to affect comprehension, as in (1), using the materials from Coulson & Williams (2005). This approach provides a powerful way to probe the relative contributions of these three systems to pragmatic processing, while avoiding the common problem of reverse inference in fMRI (Poldrack, 2006, 2011), and yielding high sensitivity and functional resolution (Nieto-Castañon & Fedorenko, 2012).

(1) She went on a fourteen-day diet, but she only lost two weeks / ounces.

The ToM regions, including the most functionally selective RH ToM region (rTPJ; Saxe & Powell, 2006), responded more strongly during jokes than during the literal control conditions (ps<0.05). In contrast, although the language and the cognitive control regions responded robustly to both jokes and literal controls relative to a low-level baseline, they did not strongly differentiate between these two conditions: a few language and cognitive control regions showed weak preferences for jokes, but if anything, these preferences were stronger in the LH regions. We therefore conclude that the RH deficits in pragmatic processing as well as the left-visual-hemifield or right-ear processing advantages for non-literal aspects of language (Coulson & Lovett, 2004; Coulson & Williams, 2005; Coulson & Wu, 2005) plausibly arise from damage to the ToM circuits rather than the RH homologues of the language regions or the RH subset of the cognitive control system.

We will discuss the implications of these findings for non-literal language more broadly. For example, some phenomena – like humor processing – may rely more heavily on the social cognition system while others (e.g., understanding metaphors) may primarily recruit the language or the cognitive control system. Furthermore, these three systems may contribute differently to pragmatic deficits in neurodevelopmental disorders (e.g., Champagne-Lavau & Joanette, 2009); for example, deficits in social reasoning may contribute to problems with non-literal processing in autism spectrum disorders, but deficits in domain-general working memory / cognitive control may contribute to pragmatic deficits in schizophrenia.
THE RELATIONSHIP BETWEEN LITERACY AND SPOKEN LANGUAGE PROCESSING
Naomi Havron and Inbal Arnon (Hebrew University of Jerusalem)
Naomi.havron@mail.huji.ac.il

Why is it harder to learn a language as an adult? Unlike children, adult learners are usually literate, and there is evidence that literacy impacts language processing (e.g. Huettig, Singh & Mishra, 2011). Literacy leads learners to attend more to individual words (lexical segmentation, Kurvers & Uri, 2006), which may hinder learning the relations between words (Arnon & Ramscar, 2012). We test the hypothesis that literacy increases lexical segmentation in spoken language by looking at 31 adults learning to read in their L2 (Hebrew) – half of them could read in their L1, while the other half was fully illiterate (all participants speak the L2). We assessed their ability to reverse pairs of spoken words in their L2 (hear: small-boy, produce: boy-small) before and after an intense L2 literacy course. This enabled two comparisons: (1) between fully illiterate and L2 illiterate at the start of the course, and (2) before and after L2 literacy course. We manipulated item frequency, to see if more frequent pairs will be harder to reverse. By looking at this unique population and comparing them to themselves after a literacy course, we are able to better isolate the effect of L1 and L2 literacy, and control variables differentiating literate and illiterate adults. We found several interesting patterns. First, frequent pairs were indeed harder to reverse, validating the processing aspect of the task. Second, L1 literacy led to better task performance at the beginning of the course, even though the task in spoken L2 (logistic regression, B=1.51, SE=.476, p < .01, controlling for L2 vocabulary and non-verbal intelligence). Third, both task performance and literacy in L2 improved after the course. Finally, L2 reading at the end of the course led to better task performance (logistic regression, B = .47, SE=.16, p < .01), but only for those illiterate in their L1. Common errors were repetition and reversal of meaning (heard: small-boy, produced: big-boy). This is, to our knowledge, the first study to follow adult illiterates before and after they learn to read. Results enhance existing findings on the effect of literacy on spoken language processing, support the notion that literacy increases lexical segmentation and undermine the claim that only age impacts word awareness (Karmiloff-Smith et al., 1996). We discuss implications for models of L2 learning.

Bibliography:
Prior research into processing VP ellipsis suggests that its interpretation involves a cost-free retrieval mechanism, since syntactically complex predicates are no more costly to retrieve than less complex ones (e.g., Frazier & Clifton, 2001, 2005; Martin & McElree, 2008). However, there is a processing penalty when an ellipsis remnant includes a “sprouted” phrase – i.e., a syntactic unit with no correlate in the previous clause, as in the sluicing sentence John ate, but I don’t know what (Frazier & Clifton, 1998; Dickey & Bunger, 2005). This penalty arguably reflects having to construct additional semantic information for the correlate at LF, e.g., John ate <SOMETHING>, but I don’t know what (Chung et al., 1995, 2011).

We examined the much-less construction in (1), which is argued to require ellipsis (Hulsey, 2008; Harris, 2013), to determine whether a similar kind of cost-free retrieval mechanism avoids complexity costs, and, if so, to what extent sprouting disrupts the retrieval process. Much less is interesting in this regard because in a sentence like John didn’t eat a cookie, much less a cake, the remnant (a cake) must form a scalar contrast with its correlate (cookie) such that eating a cake is less expected than eating a cookie. Thus negating the correlate contextually entails the negation of the remnant (see Toosarvandani, 2010, for the related let-alone construction). As such, we might expect that a sprouted adjective as in (1) would actually facilitate processing, rather than disrupt it. However, adjective sprouting inflates processing times, as previously shown for other sprouted phrases in sluicing.

To determine the remnant preferences and frequency of sprouting in general, we first examined over 1600 examples of much-less ellipsis from the Corpus of Contemporary American English (CoCA). The two most common syntactic phrase types after much less were NPs (45%) and VPs (31%). Sprouting was rare, with at most 25 instances of any type.

Experiment 1 presented 20 sentences like (2). Subjects (N=24) chose between the answers in (2a-b), either a VP or an NP with a sprouted adjective contrast. The conditions varied the presence of the domain widener any before the object NP correlate to see whether it would facilitate the scalar contrast via entailment: if the police didn’t arrest any subject, they didn’t arrest a dangerous one. We also manipulated the presence of even, which facilitates processing scalar contrast in related structures (Harris, 2013). All conditions received over 80% VP choices, but there were significantly more NP answers with any, even though it required sprouting. The results support cost-free retrieval of structure, since NPs are smaller but were not preferred; contrast with the CoCA distribution of VPs vs. NPs; and show a dispreference for sprouting.

Experiment 2, a pilot self-paced reading study (N=8), varied VP and NP remnants and whether the object correlate contained an adjective (3). Initial results for the spillover region show a higher cost for NPs with sprouted adjectives than without (d=88ms) but no effect of the adjective on VP remnants (d=2ms). This suggests that the processor constructs the scalar contrast immediately in much-less ellipsis, and that a contrastive adjective on the correlate facilitates recovery of the remnant, but does not promote expectations about the remnant type – i.e., NP over VP remnants.

Our findings support the notion of a cost-free retrieval mechanism for ellipsis that is penalized if it is necessary to augment the semantic representation, as in sprouting. This penalty persists even in much-less constructions whose interpretation might plausibly be facilitated by sprouting.

1. I’m not sure I own a hat, much less a red one. (CoCA)
2. The police didn’t (even) arrest {a/any} suspect, much less…
   a. charge one. (VP) b. a dangerous one. (NP)
3. The police | didn’t arrest | a (harmless) suspect, | much less | {charge one (VP)/a dangerous one (NP)} | which made the commissioner | very concerned.
ADAPTATION OF ACTIVE GAP FILLING IN COORDINATED SENTENCES
Akira Omaki, Aaron Apple (Johns Hopkins University)
omaki@jhu.edu

Much research has investigated syntactic priming during sentence comprehension (e.g., Traxler 2008). These studies typically test whether syntactic priming modulates biases in resolution of temporary structural ambiguities (e.g. dative alternation, main verb vs. reduced relative), but little work has investigated whether syntactic priming extends to structure building processes. To address this question, the current study investigates filler-gap dependency priming in coordinated sentences. In particular, we tested whether the active gap filling bias for early completion of filler-gap dependencies (e.g. at a verb) can be inhibited in the second conjunct as a function of the gap position in the first conjunct. Two eye-tracking during reading experiments show that the active gap filling bias can be modulated, but only when the initial filler-verb association causes a reading disruption.

**Experiment 1** (n=32) used coordinated sentences (1), in which the syntactic structure of the first conjunct is argued to prime that of the second conjunct (e.g., Frazier et al. 1984; Sturt et al. 2010). Both conjuncts included a filler-gap dependency. In the first conjunct, the gap position was manipulated to be at either the verb or the subsequent preposition, although the filler (*the manuscript*) may be temporarily analyzed as the object of the verb (*wrote*). The second conjunct involved a filler-gap dependency with a preposition gap, and the semantic fit between the filler and the verb was manipulated. If the parser actively completes the dependency at the verb, we expect a *plausibility mismatch effect* (Traxler & Pickering 1996), a reading time slowdown in the implausible filler condition (*flaws-wrote*) compared to the plausible filler condition (*review-wrote*). This 2x2 design allowed us to probe priming of filler-gap dependency completion: if the gap position in the first conjunct causes syntactic priming, plausibility mismatch effects in the second conjunct should be eliminated when the first conjunct contains a preposition gap.

For the critical verb region in the second conjunct, linear mixed effects models on the first-pass reading time demonstrated a main effect of plausibility (p=.02) but no significant interaction. This suggests that the fixation duration in the implausible conditions was slower than in the plausible conditions, regardless of the gap position in the first conjunct. Although this may suggest that active gap filling is robust to syntactic priming, the lack of priming effects may be due to the fact that the (temporary) filler-verb association in the first conjunct is plausible, and thus does not lead to reading disruption (*manuscript-wrote*).

**Experiment 2** (n=32) addressed this question by manipulating the filler-verb semantic fit in both first and second conjuncts. For the first conjunct, (2ac) used a plausible filler-verb combination with a verb gap (i), and (2bd) used an implausible filler-verb combination with a preposition gap (ii). For the second conjunct, (2ab) involved a plausible filler-verb combination (iii), and (2cd) involved an implausible filler-verb combination (iv). In contrast to Exp1, a significant interaction was found in the regression path duration on the verb in the second conjunct: the plausible first conjunct conditions showed a reliable plausibility mismatch effect (2a: 335ms vs. 2c: 451ms), while the implausible first conjunct conditions showed little fixation duration difference (2b: 372ms vs. 2d: 360ms). The lack of plausibility mismatch effect in these conditions indicates priming of the preposition gap in (2d).

These findings indicate that the preposition gap structure can in principle be primed, but only when the initial active gap filling in the first conjunct produces a reading disruption. Implications for theories of syntactic priming and adaptation will be discussed.

(1) The manuscript that the author {wrote ___ lazily | wrote lazily about ___ } was inaccurate, and the {review | flaws} that the editor wrote scathingly about ___ was critical of his writing.

| (2) | (2a) = (i) + (iii); (2b) = (ii) + (iii); (2c) = (i) + (iv); (2d) = (ii) + (iv) |
| 1st conjunct | (i) The manuscript that the author wrote ___ lazily was historically inaccurate, ... |
| 2nd conjunct | (ii) The city that the author wrote lazily about ___ was misrepresented, ... |
| (iii) ...and the review that the editor wrote ___ scathingly was critical of his writing. |
| (iv) ...and the flaws that the editor wrote scathingly about ___ were embarrassing. |
ACTIVE SEARCH MAY OVERRIDE LOCAL COHERENCE EFFECTS
Lauren Ackerman (Northwestern U), Nina Kazanina (U of Bristol) & Masaya Yoshida (Northwestern U)
lmackerman@u.northwestern.edu

The human parser is susceptible to local coherence effects, whereby the input is parsed into locally coherent portions that (sometimes) cannot be integrated into the global grammatical structures [1,2]. This study investigates the scope and strength of local coherence bias by examining how it interacts with other pressures on the parser. Consider the example in (1), in which the NP Annie’s favorite melody is locally ambiguous between the Embedded Object (EO)/Matrix Subject (MS) readings.

(1) Whenever she was trying to casually hum Annie’s favorite melody was beautiful.

Previous studies have shown that in sentences such as (1), the local coherence bias initially directs the parser to an EO analysis in the EO/MS ambiguous sentences (note that the verb hum is equibiased and occurs equally often with or without a direct object) [3]. Interestingly, however, (1) features additional pressure on the parser, triggered by the presence of a cataphoric pronoun she in the embedded clause. Namely, upon encountering the cataphoric pronoun, the parser starts an active search for its antecedent [4,5]. During the active search, the parser only seeks a potential antecedent in grammatically sanctioned positions, i.e., positions that the pronoun does not c-command [5,6]. In (1), the potential antecedent that is closest to the pronoun is Annie. However, Annie can be a licit antecedent for the pronoun only if it is part of the MS, since the pronoun in the embedded subject position c-commands the EO position. Thus, a bias toward an early and grammatical completion of the cataphoric dependency should lead the parser to analyze the ambiguous NP as the MS.

In an eye-tracking while reading experiment, we examined whether the parser’s bias toward building a locally coherent structure could be overridden by the additional pressures of an active search for the antecedent for a cataphoric pronoun. Twenty eight participants read sentences such as (2), where Gender of the pronoun (Match vs. Mismatch) and presence of Punctuation after the embedded verb (Comma vs. No Comma) were manipulated as independent factors in 2x2 factorial design.

(2)a/b. Whenever she/he was trying to casually hum Annie’s favorite melody was beautiful.
(2)c/d. Whenever she/he was trying to casually hum, Annie’s favorite melody was beautiful.

In (2a/b), the cataphor she/he should trigger the parser’s active search for the antecedent, so the parser should attempt an MS analysis of the ambiguous NP. As a result, we expect a surprisal slowdown at the name Annie in (2b) vs (2a) when it mismatches the cataphoric pronoun in gender, i.e. the so-called Gender-Mismatch Effect (GMME [2,6,7]). However, if the active search process is trumped by local coherence, the ambiguous NP in (2a) should initially be analyzed as the EO and, as a consequence, the pronoun will c-command Annie. This makes Annie an illicit candidate antecedent for the cataphoric pronoun, which should result in equal reading times at the name Annie regardless of its gender congruency with the pronoun. Conditions (2c/d), in which the MS analysis is enforced due to the comma, serve as controls. In particular, we expect a GMME at Annie in 2c/d, against which the presence or absence of a similar effect in (2a/b) will be judged.

Our results revealed a significant main effect of Gender at the name Annie on the first-pass reading times, which were significantly slower when the Gender was mismatched (p<.05). Furthermore, there was a significant interaction of Gender and Punctuation in the region following the matrix verb (i.e. at beautiful) in regression-path times, with No Comma + Mismatch read slower (p<.02). Cumulatively, these results indicate that an active search for the cataphoric antecedent initially overrides the bias toward local coherence, although local coherence effects are present at a later stage.

Folks have been studying children's moment-to-moment language processing using the visual world paradigm for over 15 years now. Much of this work has focused on the speed with which children interpret words and the degree to which children's online processing displays central properties of adult comprehension (such as incrementality, interactivity and prediction). But these methods can also be used to address core questions in language acquisition by exploring the kinds of representations that children construct during comprehension and how they change over development. I'll illustrate this with two case studies (negation and verb argument structure). In both cases, the findings suggest that children rapidly and readily construct semantic representations that are pretty abstract. Time permitting, I'll discuss how these structured semantic representations could play a central role in language acquisition.
One of the complex challenges faced by language learners is the acquisition of the syntactic structures of their native language. A rich literature suggests that a surface analysis of the speech signal could allow children to learn some aspects of the structure of their language (Morgan & Demuth, 1995). Specifically, since the prosodic structure of an utterance tends to coincide with its syntactic structure, children could exploit prosodic boundaries to detect syntactic boundaries and start building the skeleton of a syntactic structure (Christophe et al., 2008). Research examining that question has revealed that children are sensitive to this information from early on: infants can detect phonological phrases by 9 months of age (Gerken et al., 1994) and use prosodic boundaries to constrain lexical access by 13 months of age (Gout et al., 2004). While there is evidence that adults and preschoolers can use phrasal prosody to recover the syntactic structure of sentences (Millotte et al., 2008; Snedeker & Yuan, 2008), there is currently no evidence that younger children are able to do so. Here, we examined whether two-year-old French children, who are still in the process of learning the syntax of their language, exploit phrasal prosody online to constrain syntactic analysis.

Pairs of homophones belonging to different syntactic categories (nouns and verbs) were used to create locally ambiguous test sentences (e.g., a noun sentence: [la petite ferme] [est très jolie] the small farm is very nice vs. a verb sentence: [la petite] [ferme la fenêtre] the little girl closes the window, where brackets indicate phonological phrase boundaries). Although both sentences started with the same three words, they could be disambiguated by the prosodic boundary that either directly preceded the critical word ferme in the case of verbs or directly followed ferme in the case of nouns. Crucially, all words following the homophone were masked, such that prosodic cues were the only disambiguating information. Using an intermodal preferential looking procedure with an eye-tracker, two-year-olds listened to these test sentences while presented with two images displayed side-by-side on a TV screen: one associated with the noun interpretation of the ambiguous word (e.g., a farm) and the other one with the verb interpretation (e.g., a little girl closing something). Children were initially biased to look more toward the verb image (which always depicted humans). However, in the noun prosody condition, they switched their gaze toward the noun image around the end of the ambiguous word, which, taking into account saccade preparation time, suggests that they computed its syntactic category before word offset.

These results show that two-year-olds, upon hearing the first words of a sentence, exploit its prosodic structure to group words into constituents, and exploit this constituent structure to compute the syntactic category of words. Given that toddlers have access to prosodic structure well before the age of two, this suggests that such a mechanism could be active during the first steps of syntactic acquisition.

Figure 1: Proportion of looks to the noun image, relative to the beginning of the ambiguous word, in the noun and verb prosody conditions. Two-year-olds (n = 40) were initially biased to look more toward the verb image but switched to the noun image in the noun prosody condition. A nonparametric cluster-based permutation test (Maris & Oostenveld, 2007) revealed a significant difference between the noun prosody and the verb prosody conditions (p<0.01) starting about 600ms after ambiguous word onset.
Studies of development of sentence processing have focused largely on languages like English, where verb information and word order are critical for thematic role assignment. A smaller literature demonstrates that children acquiring head-final languages use case-marking to predict upcoming arguments (Choi & Trueswell, 2010; Özge et al., 2013). Yet, it has been shown that learners of head-final languages fail to use verbal morphology to detect thematic roles unless accompanied by other cues (e.g., case-marking, number of arguments) (Göksun, et al., 2008; Lidz et al., 2003; Trueswell et al., 2012). The hypothesis is that (i) verbal morphemes are acquired later than nominal markers or (ii) that case-marking is a more dominant cue for thematic role assignment as the verb generally becomes available sentence-finally in these languages.

The present study explores whether children can use verbal morphology for thematic role assignment when it is the only available cue. Specifically, we investigate whether Turkish-speaking children can use the relativizing and the passivizing morphemes to assign thematic roles to arguments that are omitted (in a sentence with no case-marking cues).

We conducted two experiments to compare the comprehension of subject- vs. object-relativizing morphemes (Exp-1: 1 vs. 2) and subject-relativizer vs. a combination of passivizer and subject-relativizer (Exp-2: 1 vs. 3). Each experiment tested three age-groups (4;0-4;6, 5;6-6;0, 8;0-9;0) using a sentence-referent matching task. Participants were presented with a picture with an animal acting on another animal and pre-recorded questions (e.g., 1-3), and were asked to answer the question by pointing to the correct animal. The correct response is the Agent for (1), and the Patient for (2) and (3). To eliminate case-marking cues, all arguments were omitted, which is acceptable in Turkish.

(1) **Hangisi it-en??**
   which push-SRel
   'Which is the one that is pushing it?'

(2) **Hangisi it-tiq-i?**
   which push-ORel-Poss.3sg
   'Which is the one that it is pushing?'

(3) **Hangisi it-il-en??**
   which push-Pass-SRel
   'Which is the one that is being pushed?'

We analyzed whether the type of the verbal morpheme had any effect on Agent Preference (i.e., # of Agent-Responses - # of Patient-Responses) in all age groups. In Exp-1, the number of Agent responses was significantly greater in the Subject-Relatives compared to the Object-Relatives only in the oldest (8-year-old) group (z=-2.54; p=.01). In Exp-2, the number of Agent responses was significantly greater in the Subject-Relatives compared to the Passive Condition in all age groups (4-year-olds: z=-3.08; p=.002; 5-year-olds: z=-2.06; p=.007; 8-year-olds: z=-2.9; p=.000).

By age 8 children could understand the subject- and the object-relativizer. The object-relativizer led to erroneous Agent Responses at younger ages, probably due to its ambiguous function (i.e., relativization, complementation). By age 4 children could differentiate the subject relativizer and a combination of passivizer and subject relativizer. The study revealed for the first time that verbal morphemes guide thematic role assignment even without any accompanying cues. Moreover, children can interpret concatenating verbal morphemes incrementally as young as 4 years of age. Thus, verbal morphemes are not necessarily acquired later than case-marking cues although case-marking cues on sentence-initial nouns might be dominating the interpretation of thematic roles during the course of online processing, causing children to ignore the late coming verbal morphemes.
Knowledge of language includes syntactic principles which appear unlearnable given the language input children receive in development. The structure dependence of auxiliary fronting in complex polar questions is a paradigm case of such a principle (Chomsky, 1980).

English complex questions are formed by fronting the main clause auxiliary (“Is the boy who is eating running?”), structure-dependent movement), not the embedded clause auxiliary (“Is the boy who eating is running?”, structure-independent movement). Since these questions are rare in child-directed speech, it has been claimed that innate constraints are needed to remove structure-independent rules from the learner’s hypothesis space (Berwick et al., 2011). Here we test whether structured meaning can be an alternative to innate syntactic constraints.

We adapted a connectionist model of language acquisition and sentence production (Chang, 2009) that learns from exposure to meaning-form pairs. Crucially, we assume that learners must be able to infer non-linguistic meanings, e.g., that the above sentence describes the two events \text{RUNNING(BOY)} + \text{EATING(BOY)}. The EATING message is given and the RUNNING message is not known by the speaker (marked as a question). Over time, the model should learn to use this message structure to acquire the correct rule.

Model input consisted of simple and complex declaratives and questions with copulas, modal auxiliaries, and progressive verbs. No complex questions with auxiliaries in both clauses occurred in training. After 500K learning episodes, the model produced a diverse set of such questions with +95% accuracy from novel meanings. Model behavior was consistent with a range of findings concerning relative clause length, structure, and auxiliary type.

Errors during development were similar to those found in children who would often repeat auxiliaries (“Is the boy that is nice is happy?”) in elicited production (Crain and Nakayama, 1987) and occasionally make embedded-auxiliary errors (“Can boys that run can jump?”) (Ambridge et al., 2008). The model suggests how children could eventually learn to retreat from these errors and corpus evidence is provided to support this explanation.

From simple questions the model learned that features in the message signalled auxiliary movement, and from complex declaratives that different parts of a sentence depended on different message components. By combining these regularities that are present in child-directed speech, it was able to generalize auxiliary movement in a structure-dependent way. This is the first explicit computational model that learns to use message structure for subject-auxiliary inversion in the production of complex questions.

References


EVIDENCE FOR A PHASE TRANSITION IN LEARNING A RECURSIVE ARTIFICIAL GRAMMAR
Whitney Tabor (University of Connecticut, Haskins Labs), Garrett Smith (University of Connecticut), & Pyeong Whan Cho (University of Connecticut)
whitney.tabor@uconn.edu

Recursive patterning is a property of most, if not all, human languages. We probed learning of a simple recursive pattern (similar to $a^nb^n$) in an artificial grammar task. Such learning is nontrivial because it involves generalizing from a finite sample to an infinite principle. We find evidence that, instead of generalizing immediately to deeper levels of embedding, learners exposed to progressively deeper levels of recursion fail to generalize at first, then abruptly shift to generalizing when the depth becomes sufficiently great (see [2]). Yet, even after the jump, they do not generalize perfectly and continue to show improvement with training. Such graded, qualitative shifts are puzzling under discrete grammar models. Dynamical systems theory offers insight via the notion of phase transition: a continuous change of a parameter (called a control parameter) results in a qualitative change in the organization of the system (e.g., when temperature decreases over a suitable range, matter changes from liquid to solid). Here, we provide evidence that depth of recursion encountered acts like an incremental control parameter. We present both human artificial grammar learning data and an artificial neural network in which continuous change of the weight parameters results in a bifurcation (the standard mathematical model of phase transition) associated with a switch from finite-state to recursive grammar structure.

When undergraduate participants clicked a computer mouse, one of four boxes arranged in a diamond on screen changed color. The task was to click the box that would change color next (Locus Prediction task). A context-free grammar determined the order of changes: $S \rightarrow 1S234$, $S \rightarrow \epsilon$, where 1, 2, 3, and 4 are labels (not visible) for the four boxes. This grammar was used to generate level 1 sentences (L1s = 1234), L2s (11234234), etc., through L5s, strung end to end. There were 756 trials (color changes) divided into five stages. Table 1 lists the training schedule by condition. Note the jumps in conditions B (L3 to L5), C (L2 to L4), and D (L2 to L5).

<table>
<thead>
<tr>
<th>Stage:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>L1</td>
<td>L1-2</td>
<td>L1-3</td>
<td>L1-4</td>
<td>L1-5</td>
</tr>
<tr>
<td>B</td>
<td>L1</td>
<td>L1-2</td>
<td>L1-3</td>
<td>L1-3</td>
<td>L1-5</td>
</tr>
<tr>
<td>C</td>
<td>L1</td>
<td>L1-2</td>
<td>L1-2</td>
<td>L1-4</td>
<td>L1-5</td>
</tr>
<tr>
<td>D</td>
<td>L1</td>
<td>L1-2</td>
<td>L1-2</td>
<td>L1-2</td>
<td>L1-5</td>
</tr>
</tbody>
</table>

A generalized linear mixed effects model (GLMM) (logit linking function, reference level = condition A) was fit to participants’ prediction accuracy on 4 to 1 and 4 to 2 transitions in the first L5 sentence, which required generalizing to deeper embeddings than in training. A likelihood ratio test revealed that condition significantly affected participants’ recursive generalization ($\chi^2(3) = 44.134, p < .001$). Post-hoc pairwise comparisons showed training on L3s (conditions A, B, and C) led to significantly better generalization (all $ps < .001$) than training only up to L2 (condition D). Thus, there is a sharp increase in performance after training on sentences of depth 3. This does not appear to be a discrete jump to a perfectly recursive grammar, though. The GLMM above suggests that there is continued improvement after the jump at L3: Condition A (L1–4) generalized marginally better than condition B (L1–3 only) ($p = .052$). We believe this type of graded, qualitative shift is best explained by a phase transition in participants’ learning of the grammar implicit in the training.

In addition, we built a neural network in which inputs specify recurrent hidden unit dynamics, which are mapped to outputs by linear separators (similar to [1]). Through manual exploration of the weight space, we discovered a continuous weight-change trajectory that shifts the model from (finite state) processing of only one and two levels of embedding to (infinite-state) processing of arbitrary levels via a bifurcation, improving gradually, similar to the human participants. Graded emergence of recursion is possible because the model encodes embeddings with a fractal, which can morph continuously. We believe the bifurcation occurs at depth 3 because the chances of discovering a non-recursive solution in the weight space go down rapidly as the evidence for recursion mounts. We suggest that dynamics on fractals offers a viable alternative to traditional grammar models, suited to graded psycholinguistic phenomena.

BREAKFAST WORKSHOP FOR STUDENTS AND EARLY CAREER RESEARCHERS

GETTING YOUR PAPER PUBLISHED:
WHAT YOU NEED TO KNOW ABOUT PUBLICATION

Adam Fraser, Associate Publisher, Journal of Memory and Language
(a.fraser@elsevier.com)

Academic publishing is a competitive field. It can be daunting for researchers starting their careers. But there are important steps that, as an Author, you can follow in preparing your manuscript for a successful publication. Hear the publisher’s perspective.
SNAP JUDGMENTS: SMALL N ACCEPTABILITY PARADIGM (SNAP) FOR LINGUISTIC JUDGMENTS

Kyle Mahowald (MIT), Peter Graff (Intel), Jeremy Hartman (UMass), & Edward Gibson (MIT)
kylemaho@mit.edu

Recently, the practice of researchers’ relying on intuition in syntax and semantics judgments has been questioned due to concerns that these judgments are potentially erroneous and not quantitatively rigorous (Wasow & Arnold 2005; Myers 2009; Gibson & Fedorenko 2010; Gibson, et al., 2013). With crowdsourcing platforms like Amazon’s Mechanical Turk, it is possible to perform large-scale experiments and obtain results from reliable workers in a matter of a few hours. Still, some researchers have noted that it seems like overkill to run a large-scale experiment for a seemingly obvious contrast (“Colorless green ideas sleep furiously” vs. “Furiously sleep ideas green colorless.”) Here, we propose SNAP (Small N Acceptability Paradigm) Judgments, which make obtaining formal linguistic judgments easier and cheaper, while still allowing statistically valid conclusions to be drawn.

**Replication of Sprouse, Schütze, & Almeida (2013):** First, we replicated the findings of Sprouse et al., who performed formal experiments for 150 contrasts sampled from Linguistic Inquiry from 2001-2010. We formally tested another 100 contrasts from the same LI data set which were not tested by Sprouse et al., writing 12 unique items for each. In a forced choice study (asking participants to choose whether Sentence A sounds more natural than Sentence B), 92/100 contrasts showed significant effects in the predicted direction, while 6 showed trends in the direction opposite of what was predicted (2 significantly). These results are within the range found by Sprouse et al., but this is arguably not an acceptable error rate (Gibson et al., 2013).

**SNAP Judgments:** Given the distribution over effect sizes from the forced choice experiment (ranging from .5 to .99), we can compute how many people one should ask in order to be 95% confident that the results of a full experiment would have an effect size over .9 with a lower 95% bound over .75. We do this in a Bayesian framework by placing a beta prior (empirically fit, as shown in Fig. 1 to be Beta(.6, .6)) over all experiments, which we can combine with the results of the binomial experiment (i.e., how many people prefer Sentence A out of the total number surveyed) to get a posterior estimate. The Beta(.6, .6) prior means that, if we ask n people and all n prefer Sentence A, we can act as if n + .6 preferred Sentence A and .6 preferred Sentence B. In order to achieve the confidence thresholds above, we found that we have to ask 7 people. This result is robust to several plausible ways of setting the prior.

**Testing SNAP:** We provide empirical support for this recommendation by randomly sampling 7 data points (unique subjects and items) from our experiment, and—of those that are unanimous—computing the results in the full experiment. In 100 simulations, when the sample was unanimous, the mean preferring Sentence A was .93 with a 95% CI of [.79, .99]—consistent with the prediction made using the prior. The lowest value obtained in any simulation was .72, still a strong effect. We thus make the following recommendation:

> When you have a strong intuition that Sentence A is better than Sentence B, present 7 unique subjects each with a unique instance of the contrast between Sentence A and B. If the result is unanimous, you can conclude that the most likely effect size in a large experiment is .93 with a 95% CI of [.75, 1]. If all participants are not in agreement, run a formal experiment to find out the source of the variation.

**Conclusion:** As demonstrated empirically, SNAP Judgments allow for robust, statistically valid generalizations with minimal cost in time and money.
Recent evidence suggests that cognitive pressures associated with language acquisition and use could affect the organization of the lexicon. On one hand, consistent with noisy channel models of language (e.g., Levy 2008), the phonological distance between wordforms should be maximized to avoid perceptual confusability (a pressure for sparsity). On the other hand, a lexicon with high phonetic regularity would be simpler to learn, remember and produce (e.g., Monaghan et al., 2011) (a pressure for clumpiness). Here we investigate wordform similarity in the lexicon, using measures of word distance (e.g., phonological neighborhood density) to ask whether there is evidence for sparsity or clumpiness of wordforms in the lexicon. We develop a novel method to compare observed lexica to phonotactically-controlled baselines that provide a null hypothesis for how clumpy or sparse wordforms would be as the result of only phonotactics. Beyond this phonotactic baseline, we present evidence for widespread clumpiness, using a highly-controlled comparison of four languages, as well as a broad coverage analysis of 115 languages.

1) Simulating phonotactically-controlled lexica: We first studied monomorphemes in Dutch, English, German (CELEX) and French (Lexique). To control for phonotactics, we built several generative models of lexica: ngrams over phones, ngrams over syllables, and a PCFG over syllables. After training, we evaluated each model on a held-out dataset to determine which generative process most accurately captured each language. We then used the best model to generate simulated lexica for which there is no pressure for sparsity or clumpiness, beyond phonotactics. Comparison of the wordform distance in the real lexicon (a test statistic) to these phonotactically-controlled lexica (providing a null distribution of the test statistic) allows us to determine if the real lexicon is significantly sparser or clumpier than would be expected by its phonotactics alone. Across a number of measures of lexical clustering (e.g. string edit distance, clustering coefficients, etc.) we found that the real lexicon had a tendency to be clumpier than the best phonotactically-controlled lexica for most measures. For instance, in all languages the average edit distance between two words in the real lexicon was smaller than the edit distance in the simulated one (all ps < .05). Note that this is not the result of morphological regularities since only monomorphemes were used.

2) Cross-linguistic frequency: A second analysis shows that a pressure for clumpiness can also be found in the frequency patterns of word use (see Frauenfelder et al. 1993). If the lexicon is influenced by a pressure for clumpiness, we should expect that higher-frequency words exhibit stronger patterns since they are used more often and thus subject to more communicative or cognitive optimization. Using orthographic lexicons from 115 languages extracted from Wikipedia (top 20,000 most frequent wordforms in each language and treating each length separately), we found that almost all languages (111/115), show significant (p < .05) positive Spearman correlations between log-frequency and the number of neighbors (mean r = .18). This effect is, in part, driven by a correlation in 112/115 languages for more orthographically probable words (as measured by an ngram model over letters) to be more frequent (mean r = .22). Thus, frequent wordforms tend to be phonotactically likely and have more neighbors—suggesting a pressure for clumpiness that becomes stronger with use.

3) Semantic clumping: For each pair of words in English and in French, we computed the pair’s semantic similarity using LSA on Wikipedia and matched it to a pair of words in each simulated lexicon. Pairs with smaller edit distance are more likely to be semantically similar than phonologically distinct words (ps < .001) at a rate above chance (ps < .05). The trend for semantically similar words to be phonologically similar is evidence for lexical clustering.

Conclusion: Across a variety of measures, we showed that lexica have the tendency to be clumpier than expected by chance. This reveals a fundamental drive for regularity in the lexicon that conflicts with the pressure for words to be as phonetically distinct as possible.
Research on syntactic priming has focused mostly on mono- or bilingual processing. So far, only few studies have examined structural priming in closely related varieties [1, 2]. These bivarietal studies found stronger priming within than between language varieties and suggest separate lemma representations of cognates across varieties.

For example, Vorweg et al. [1] examined priming of pre- and post-nominal attributes within Bernese German, a Swiss German dialect, as well as between Bernese German and Standard German. When priming within Bernese German, effects were rather high (same nouns: 40%, semantically related nouns: 23%, different nouns: 8%) and comparable to findings of other monolingual studies using similar experiments [3, 4]. In contrast, when priming from Bernese German to Standard German or the other way round, effects were very low (same nouns 4%/9%, semantically related nouns: 4%/6%, different nouns: 3%/2%).

These results imply that Bernese German and Standard German are represented separately at the lemma level, possibly like an L1 and an L2. However, the effects between varieties are surprisingly small, even for separately represented languages, and there might be other factors influencing these effects. One possible explanation is that priming did not work due to the two varieties’ slightly different use of relative pronouns in the post-nominal construction (Bernese German: indeclinable relative pronoun wo; Standard German: declinable, gender-marked pronoun der/die/das). If the different use of pronouns caused the low priming effects between Bernese German and Standard German, we should expect higher effects when priming within Standard German, as the same pronoun would appear in prime and target. However, the low effect strength might also be due to an inhibition in Bernese speakers given the sociolinguistic status of Standard German in German-speaking Switzerland. It is often perceived as a distant, very formal and sometimes even foreign language for native Swiss German speakers and they might therefore show little alignment. If this is the case, we should expect higher priming between Bernese German and another dialect, which is clearly different from Bernese German but has a similar, positive status.

In Experiment 1, we tested priming of pre- and post-nominal attributes from Standard German to Standard German in native Bernese German speakers. Contrary to the assumption that there could be higher effects within Standard German than between varieties, the results showed virtually no priming (same nouns: 5%, translation equivalents and different nouns: 0%). To compare, an overall priming effect of 14% was found when priming from L2 to L2 (English) in Dutch speakers [5]. It is therefore unlikely that the different use of relative pronouns caused the weak effects in between-variety priming. Rather, we assume the reason for low priming to be the sociolinguistic distance.

In Experiment 2, we investigated priming of the same noun-phrase structures from Valais German to Bernese German in native speakers of Bernese German, in order to test priming between varieties that have a similar sociolinguistic status. Valais German and Bernese German are both regional dialects and are rated among the most popular dialects in Switzerland. However, there is also a rather large distance between the two, and Bernese speakers sometimes find Valais German more difficult to understand than Standard German and they usually cannot speak this dialect. In contrast to the low priming effects between Standard German and Bernese German, this experiment revealed considerable priming effects (same nouns: 18%, translation equivalents: 3%, different nouns: 1%).

To conclude, we found priming effects between two dialects with similar status. The effects were smaller than priming within one dialect, suggesting a separate mental representation. The nonexistence of priming effects within Standard German further supports the separate account and implies an inhibition due to the sociolinguistic status of this variety.

LEXICAL (MIS)ALIGNMENT: AUTOMATIC OR STRATEGIC BEHAVIOR?
Anouschka Foltz (Bielefeld University), Judith Gaspers (Bielefeld University), Kristina Thiele (University of Cologne), Philipp Cimiano (Bielefeld University) & Prisca Stenneken
(University of Cologne)
anouschka.foltz@uni-bielefeld.de

Alignment is pervasive in communication. Speakers tend to adopt aspects of their interlocutors’ linguistic and non-linguistic behavior. This study focuses on lexical alignment, where speakers adopt lexical items used previously by their interlocutors. Current theories of (lexical) alignment tend to focus on either automatic or strategic processes. For example, Pickering & Garrod’s (2004) interactive alignment model assumes that lexical alignment is based on low-level priming and thus largely automatic. Brennan & Clark (1996), on the other hand, assume that lexical alignment is partner-specific (based on conceptual pacts with a particular partner) and thus largely strategic. This study uses a triadic conversational situation to explore the extent to which lexical alignment is strategic and/or automatic.

Forty pairs of adult native German-speakers played a picture-placement game. During each round, one participant told his or her partner where to place three different objects in a 3x3 grid. After each round, participants switched roles. Information about the order in which the objects had to be placed was given by a voice over headphones (i.e. a third entity not physically present in the conversation). For critical trials, participants had to name an object previously named by their interlocutor. We manipulated three factors: Pictures were either line drawings of common objects that could be given two different names (e.g. pitcher/jug) or tangram shapes that could be interpreted in different ways (e.g. arrow/rocket). The interlocutor had either named the object one (short lag) or three (long lag) rounds previously (e.g. pitcher). And the voice over headphones either provided an alternative name for the object (e.g. jug; prime) or provided no name (e.g. the third object; no prime) immediately before a critical trial. If lexical alignment is largely automatic, participants should use the name provided by the voice over headphones in the prime conditions. Furthermore, lag (remembering the interlocutor’s object name) and picture type (higher risk of misunderstanding for tangram shapes than line drawings) should not affect alignment. But if lexical alignment is largely strategic, participants should use the name provided by their interlocutor in both the prime and no prime conditions. In addition, alignment should be stronger for short lags and tangram shapes compared to long lags and line drawings.

(1) Proportions of lexical alignment

The results are shown in Figure (1). A mixed-logit analysis showed a strong effect of prime (estimate = -3.7056; z = -8.692; p < .0001): Participants aligned with their interlocutor when the voice over headphones provided no alternative name. When the voice provided an alternative name, participants typically used it, even though it mismatched the name their interlocutor provided. A main effect of picture type (estimate = 0.3956; z = 2.416; p < .05) revealed stronger alignment for tangram shapes compared to line drawings. Exploration of a reliable prime x lag interaction (estimate = -0.4052; z = -2.440; p < .05) revealed a reliable effect of lag for line drawings in the no prime conditions (estimate = 0.4064, z = 2.010, p < .05).

The results suggest that both automatic and strategic processes are involved in the lexical alignment process and highlight the need for theories of alignment in which both automatic and strategic processes play a role. We propose that alignment has a strong automatic component, but that strategic behavior is not merely a repair mechanism in cases of miscommunication. Instead, interlocutors are sensitive towards aspects of the situation, such as the difficulty of the task, the risk of misunderstanding etc. and can “decide” when the task requires more strategic behavior.
INPUT VARIABILITY AND THE SPECIFICITY OF PERCEPTUAL RECALIBRATION

Eva Reinisch (Ludwig Maximilian University Munich) & Holger Mitterer (Università ta’ Malta)
evarei@phonetik.uni-muenchen.de

Listeners use lexical or visual context to recalibrate auditory speech perception. After hearing an ambiguous auditory stimulus between /p/ and /t/ in unambiguous lexical context (e.g., keep[p/t] -> /p/, meep[p/t] -> /t/) or visual context (i.e., presence vs. absence of lip closure) an ambiguous auditory-only stimulus is perceived in line with the previously experienced lexical or visual stimuli. However, it remains unclear how specific such learning is. Kraljicic and Samuel (2006, PBR, 13, 262-268) used lexical context to guide recalibration and found quite general learning: Recalibration following learning about the voiced-voiceless distinction in stops generalized during test from alveolar stops (/d/-/t/) to labial stops (/b/-/p/). But Reinisch et al. (2014, JPhon, 45, 91-105) using visual context did not find generalization from a labial-alveolar stop contrast (/b/-/d/) to a similar contrast that differed in manner of articulation (/m/-/n/). This raises two questions: first, do lexical and visual context indeed trigger the same recalibration process (as suggested by van Linden & Vroomen, 2007, JEP:HPP, 33, 1483-1494), and second to what extent does perceptual recalibration depend on the specific phoneme contrast involved.

Regarding the first question, in the typical visual-context paradigm, the input usually remains quite uniform. That is, participants hear only critical items like /aba/ and /ada/ repeated multiple times. In contrast, in the lexical-context paradigm, lists of 10-20 critical items are presented intermixed with fillers. The first experiment of the present study tested whether recalibration by visual-context also occurs with more varied input, similar to typical lexical-context experiments. During exposure two groups of listeners watched video clips of a speaker producing words and nonwords and performed a lexical decision task. Critical items were minimal word pairs differing in /p/ or /t/ for which the audio signal was ambiguous between the two sounds. However, one group of participants saw the speaker produce /p/ whereas the other group saw the speaker produce /t/. Following this exposure, participants performed a phonetic categorization task categorizing sounds along an audio-only [a:p]-[a:t] continuum. Results showed that listeners had recalibrated their /p/-/t/ category boundary. Participants who had seen the speaker produce /p/ (i.e., lip closure) on critical words during exposure categorized more sounds along the [a:p]-[a:t] continuum as [a:p] than participants who had seen the speaker produce /t/. This shows that recalibration based on visual-context information is possible even with more stimulus variation than in the typical visual-context recalibration experiments.

To answer the question about the units of recalibration as indicated through generalization across phoneme contrasts, an analogous experiment was run with lexical disambiguation of the ambiguous sounds. That is, the set of critical items during exposure was exchanged for words that indicated whether the ambiguous sound should be /p/ or /t/ (e.g., a sound between /t/ and /p/, at the end of keep[p/t] must be /p/). Disambiguation was again manipulated as a between-participant factor. Importantly, during test participants now not only categorized an [a:p]-[a:t] continuum but also an [a:m]-[a:n] continuum to test generalization (blocked, order counterbalanced). Results showed the typical perceptual recalibration effect: if the ambiguous sound was heard in words like keep[p/t] during exposure, more tokens along the continuum were interpreted as [a:p]. However, in line with the visual-context results by Reinisch et al. no generalization to the nasal labial-alveolar contrast was found (i.e., no group difference in the categorization of the [a:m]-[a:n] continuum).

The experiments of the present study hence support two conclusions: Both, visual-context and lexical-context recalibration work similarly, however, subtle differences like differences in the decay rate of the effects may warrant further investigation. Secondly, for both types of contexts recalibration appears rather specific. The finding that recalibration of the labial-alveolar place of articulation contrast does not generalize across manner of articulation contrasts with theories that assume that the objects of speech perception are abstract features or intended gestures.
Lexically driven perceptual learning is a powerful mechanism that tailors speech perception and recognition to the ambient environment by identifying and adjusting for patterns of variation in incoming speech [1]. Previous research has investigated the nature of these adjustments by testing whether perceptual learning of segmental variation based on exposure to a given talker (e.g., a talker with an atypical fricative or stop consonant contrast) is applied talker-specifically or whether adjustments transfer to new talkers who produce similar patterns of variation. At present, evidence for cross-talker transfer is mixed [2,3]. One hypothesis presented by Kraljic and Samuel [2] is that adjustments are talker-specific when the phonetic cues being learned provide talker indexical information, as in the spectrally-cued fricative contrasts they tested, whereas learning is talker-independent when the phonetic cues being learned provide weaker cues to talker identity, such as for the temporally-cued stop consonant contrasts they tested.

Here, two experiments investigated potential cross-talker transfer of learning following exposure to a talker with a novel system of vowel variants. Vowel variation has received little attention in this research area (though see [4]), despite the fact that vowel variation abounds in English, requiring listeners to routinely accommodate such variation. Further, vowel productions convey rich talker indexical information, thus providing a further test case for Kraljic and Samuel’s [2] hypothesis about constraints on cross-talker transfer of learning.

The experiment design involved a between subjects modification of Maye et al.’s exposure-test vowel adaptation paradigm [5]. For the exposure phase, participants passively listened to a 20-minute story spoken by a female with a novel English accent, characterized by a clockwise back vowel lowered chain shift of the vowels /u, o, a/ (e.g., /u/ as in goose sounded more like [u] as in good; /o/ sounded more like [ɔ] as in goat, etc.). The control group listened to the same talker read the same story but in a standard-sounding American English accent. The test phase was an auditory lexical decision task containing back vowel lowered word forms as test items along with filler words and nonwords. In Experiment 1 (N = 27), half of the critical items were produced by the trained female and half by a new but qualitatively similar-sounding female (presentation blocked by talker). Experiment 2 (N = 30) was identical except the new talker was a qualitatively distinct male. Greater endorsement of vowel-shifted forms (i.e., ‘word’ response) in the experimental versus control condition is taken as evidence of perceptual learning.

Lexical decisions were analyzed using mixed logit regression with the maximal random effect structure justified by the design. In both experiments, participants were significantly more likely to endorse the trained talker’s back vowel lowered pronunciations following exposure to this novel chain shift than following exposure to the same talker with standard-sounding vowels (βs > 0.36, zs > 2.8, ps < .01), indicating perceptual learning of the novel vowel system. Further, participants endorsed significantly more back vowel lowered forms from both the new female talker (Exp1) and the new male talker (Exp2) in the experimental versus control exposure conditions (βs > 0.25, zs > 2.2, ps < .05), indicating cross-talker transfer of learning. The finding that learning transferred to the qualitatively different male rules out the possibility that cross-talker transfer was driven by perceptual voice similarity. Together, these results suggest that high variability exposure (e.g., a 20-minute story containing several hundred vowel-shift instances) can induce talker-independent learning, even when the phonetic cues being learned are highly indexical (cf. the hypothesis in [2]).

THE ROLE OF SPEECH PRODUCTION IN PHONOLOGICAL DECODING DURING VISUAL WORD RECOGNITION: THE CASE OF PHONOTACTIC REPAIR

Yue Sun¹, Sharon Peperkamp¹ (École Normale Supérieure, Paris, France)
yue.sun@ens.fr

Psycholinguistic models of reading propose that, during visual word recognition, readers rely not only on word’s orthography, but also on the phonological code that is converted from print [1]. Several neuroimaging studies suggest that the speech production system is involved in the phonological processing during reading [2]. However, these are correlational studies that do not allow us to determine a causal involvement of the speech production system. In the present study, we directly investigate the role of this system in phonological processing during reading.

We focus on a case of phonotactic repair during reading. In French, words cannot begin with the sequence /tl/, and French listeners perceptually repair auditorily presented illegal non-words such as /tlamo/ into legal /klamo/ [3]. Importantly, this perceptual repair is not limited to the auditory modality but also occurs when illegal sequences are presented visually [4]. Here, we show that the effect is driven by activation of the speech production system.

We used a visual masked priming paradigm (Fig 1). French participants performed a lexical decision task on written /kl/-initial target words (e.g., clavier ‘keyboard’) that were preceded by a written prime that was either /pl/-initial (legal, e.g., plavier), /tl/-initial (illegal, e.g., tlavier), or unrelated (e.g., grière). The priming effects for plavier and tlavier were defined as the decrease in reaction time compared to the baseline unrelated prime. As plavier and tlavier are neighbors of the target word clavier, they should both yield a priming effect. Moreover, the /tl/-/kl/- phonotactic repair should make tlavier be phonologically identical to clavier and thus result in an enhanced priming effect. For half of the participants, the lexical decision task was combined with articulatory suppression (AS) which occupied their articulatory system with repetitive production of a CV syllable. An ANOVA with the factors Prime (plavier vs. tlavier) and Condition (with vs. without AS) revealed an interaction (F(1, 46) = 6, p < 0.05), which was due to the fact that, as predicted, without AS the priming effect was larger for tlavier than for plavier (F(1, 46) = 10, p < 0.01), while no such difference was obtained with AS (F < 1) (Fig 2). The absence of a main effect of AS, furthermore, shows that AS does not interfere with the visual processing of printed letters, indicating a role of the articulatory system in their phonological processing only.

Together, our results suggest that phonological mediation during reading is achieved by covert articulation of a word’s phonological code, and that reading depends on an interactive network that involves both the language perception and production modules.

Fig 1: Visual masked priming paradigm with lexical decision on the upper case target.
Fig 2: Priming effect across Prime (plavier vs tlavier) and Conditions (with vs without AS)

The Mismatch Negativity (MMN) paradigm has been used to investigate whether some speech sounds are stored with a less specific encoding than others (Lahiri and Reetz, 2002; Scharinger et al. 2012; among others), as suggested by theories of lexical representations like the Featurally Underspecified Lexicon (FUL; cf. Lahiri and Reetz, 2002; 2010). FUL suggests the coronal place of articulation is a default or underspecified feature at the lexical representational level, a proposal consistent with evidence from formal phonology. FUL also makes language-specific predictions, as in the case of vowel height in English, in which phonetically specified mid vowels are encoded in the linguistic system without any height information. Previous neurolinguistic literature has largely confirmed these claims with stop consonants and vowels.

We tested coronal underspecification with the fricatives /s/ and /f/. Fricatives provide a test-case similar to vowels (unlike stop consonants), which can be easily presented in isolation. If coronal is an underspecified feature, then a deviant coronal sound should elicit a strong MMN when labial sounds are used as standards, but not vice versa. In addition to /s/ and /f/, we also use this paradigm as a diagnosis for the representation of /h/, a subject of theoretical debate. Some researchers consider /h/ to be another type of default sound (Goldsmith, 1981), whereas FUL suggests /h/ may be specified for a laryngeal place of articulation (Lahiri and Reetz, 2010). Therefore we compared /s/ (predicted to be underspecified for place of articulation) and /h/.

24 participants were tested using a passive oddball technique with four blocks (150 deviants, 850 standards): deviant [f] in standard /s/, [s] in /f/, [h] in /s/, and [s] in /h/. In the deviant [f] in standard /s/ block, subjects hear 3-8 examples of /s/ before each /f/ token. Measurements were taken at Cz for the 150-250ms time window, though they also hold for the other expected fronto-central sites. In the /f/ and /s/ comparisons, an interaction is detected between phoneme and deviance (Interaction $F=11.96(1,23), p<.005$). A strong MMN is detected for /s/, but not for /f/, following the predictions of the FUL model. There is no strong difference, however, between the /s/ and /h/ comparisons (Interaction $F(1,23)=.015(23), p=.9$). See figure 1 below.

The first important finding we report is that the predictions of FUL are borne out in the /s/ and /f/ contrast. The second and perhaps most interesting finding is that, according to the neurolinguistic evidence presented here, /h/ does not seem to have place of articulation on a par with that of /f/ since the two do not pattern together in inducing a strong asymmetric MMN response when compared to an underspecified sound (/s/). This study then supports a revised FUL model where the place properties of /h/ are unspecified and provides another case study of how the use of neurophysiological evidence can be used to advance debates between competing linguistic theories.

Figure 1: Difference waves of the /f/ and /s/ contrasts (left) and /h/ and /s/ contrasts (right).
The architectures and mechanisms for bilingual language processing are more complicated than those of monolingual language processing because a bilingual has phonemes, graphemes, words, syntactic rules, and so on from multiple languages to choose from. How does this additional choice affect language processing? I first present evidence that bilingual speakers, readers, and listeners activate representations of the language that is not relevant at the moment. Next, I discuss possible mechanisms bilinguals might use to “zoom in” on only the relevant language, with a consideration of both linguistic cues (e.g., the talker’s native language), and extralinguistic language cues (e.g., the talker’s face). I will briefly reflect on the debate of whether language selection and/or switching processes genuinely convey cognitive advantages to bilinguals. In my conclusions, I will argue that we need to develop more detailed and explicit accounts of bilingual language processing.
Human language is hierarchically structured, and mental representations of such structure are necessary for successful language processing. In speech, however, hierarchical linguistic structures, such as words, phrases, and sentences, are not clearly defined physically and must therefore be internally constructed during comprehension. How multiple levels of abstract linguistic structure are built and concurrently represented remains unclear. We demonstrate that, during listening to connected speech, cortical activity of different time scales is entrained concurrently to track the time course of linguistic structures at different hierarchical levels. Critically, entrainment to hierarchical linguistic structures is dissociated from the neural encoding of acoustic cues and from processing the predictability of incoming words. The results demonstrate syntax-driven, internal construction of hierarchical linguistic structure via entrainment of hierarchical cortical dynamics.
Sentences with global structural ambiguities have been shown to be easier to process than corresponding disambiguated sentences (Traxler et al. 1998, van Gompel et al. 2000, 2001, 2005). These data have been used as evidence against constraint-based models of syntactic parsing, which predict that ambiguity causes competition and processing difficulty. Instead, they support a variable-choice reanalysis model of sentence processing, which posits that the parser variably chooses one of the possible structures, leading to difficulty due to reanalysis when the sentence is disambiguated toward any one structure. Although there is little evidence for competition in syntactic attachment (Clifton & Staub, 2008), competition effects have been reported for pronouns that have multiple potential referents. This has been used to argue for constraint-based models of pronoun resolution (Badecker & Staub 2002). Here we present evidence from eye movements during reading that both PP attachment and pronoun reference show an ambiguity advantage in reading. This finding suggests that assignment of pronominal reference is more like structural attachment than previously believed. We argue that these findings are best explained by extending variable-choice reanalysis models to include pronominal reference.

In the eye movement study \(N_{\text{sub}}=57\), we compared ambiguous (AMBIG) and unambiguous (HIGH/LOW) PP attachment and pronoun reference (see (1)). On the critical region 4, reading times (by-subject means in the table) on all measures were longer for pronoun conditions, likely due to length differences. First Pass Time showed a numerical advantage for low conditions as compared to high \(t = 1.83\) and a significant advantage for ambiguous conditions \(t = -2.96\). Go-Past Time also showed an ambiguity advantage \(t = -2.89\). Crucially, we found no significant interactions between ambiguity type and ambiguity or height. The results for the PP-attachment conditions are congruent with previous studies (van Gompel et al. 2001). The surprising result is that pronominal reference patterns similarly to attachment, showing an ambiguity advantage in Go Past Time. A second experiment using an online ambiguity detection task \(N_{\text{sub}}=40\) confirms that comprehenders perceive the ambiguous pronoun and PP attachment conditions as equally ambiguous.

This qualitative similarity suggests that pronoun reference, like attachment, is initially variably assigned, giving rise to a reanalysis-based cost for disambiguation. We also find a numerical advantage for low conditions, which we interpret as a recency advantage (e.g. Lewis & Vasishth, 2005). This extension of the variable-choice model critically requires that gender cues do not circumvent the initial variability of reference assignment. However, the results of our second experiment suggest that subjects can use gender cues to identify referential ambiguity. We argue that this contradictory result represents 'strategic' processing, and note that there is independent evidence supporting a failure to use gender cues in a task targeting 'automatic' processing (Greene et al. 1992), as with the natural reading in our first study. We also note that although the stereotypical gender manipulation had the predicted effect on reading times, it was less effective at explicit disambiguation.

\begin{verbatim}
1 saw/ [AMBIG] the brother/ of the waiter/ ...
   [HIGH] the brother/ of the waitress/ ...
   [LOW] the sister/ of the waiter/ ...
   ...[PP] with a beard/ but we didn’t talk for long.
   ...[Pronoun] when he visited the restaurant/ but we didn’t talk for long.
\end{verbatim}

<table>
<thead>
<tr>
<th>PP / Pron.</th>
<th>First Pass Time</th>
<th>Go-Past Time</th>
<th>Total Time</th>
<th>% Judged ambiguous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambig</td>
<td>495</td>
<td>797</td>
<td>961</td>
<td>60%</td>
</tr>
<tr>
<td>High</td>
<td>561</td>
<td>920</td>
<td>1056</td>
<td>66%</td>
</tr>
<tr>
<td>Low</td>
<td>528</td>
<td>856</td>
<td>938</td>
<td>57%</td>
</tr>
<tr>
<td>Ambig</td>
<td>801</td>
<td>975</td>
<td>1305</td>
<td>53%</td>
</tr>
<tr>
<td>High</td>
<td>881</td>
<td>1132</td>
<td>1426</td>
<td>35%</td>
</tr>
<tr>
<td>Low</td>
<td>822</td>
<td>1087</td>
<td>1384</td>
<td>36%</td>
</tr>
</tbody>
</table>

SATURDAY 11:00
A central claim of embodied theories of cognition is that sensory representations are routinely activated and influence language processing even in the absence of relevant sensory input (cf. Pulvermüller, 2005; Wassenburg & Zwaan, 2010). We tested the influence of color representations during language processing in three visual world eye tracking experiments. The method is particularly well suited to investigate this issue because the availability of relevant visual input can be manipulated.

We made use of the phenomena that when participants hear a word that refers to a visual object or printed word they quickly direct their eye gaze to objects or printed words which are similar (e.g. semantically or visually) to the heard word. We used a look and listen task which previously has been shown to be sensitive to such relationships between spoken words and visual items. In Experiment 1, on experimental trials, participants listened to sentences containing a critical target word associated with a prototypical color (e.g. ‘...spinach...’) as they inspected a visual display with four words printed in black font. One of the four printed words was associated with the same prototypical color (e.g. green) as the spoken target word (e.g. FROG). On experimental trials, the spoken target word did not have a printed word counterpart (SPINACH was not present in the display). In filler trials (70% of trials) the target was present in the display and attracted significantly more overt attention than the unrelated distractors. In experimental trials color competitors were not looked at more than the distractors. In Experiment 2 the printed words were replaced with line drawings of the objects. In order to direct the attentional focus of our participants toward color features we used a within-participants counter-balanced design and alternated color and greyscale trials randomly throughout the experiment. Therefore, on one trial our participants heard a word such as ‘spinach’ and saw a frog (colored in green) in the visual display. On the next trial however they saw a banana (in greyscale) on hearing ‘canary’ (bananas and canaries are typically yellow), etc. The presence (or absence) of color was thus a salient property of the experiment. Participants looked more at color competitors than unrelated distractors on hearing the target word in the color trials but not in the greyscale trials, i.e. on hearing ‘spinach’ they looked at the green frog but not the greyscale frog. Experiment 3 was identical to Experiment 2, except that the visual display was removed at the sentence onset, after a longer preview. This experiment examined whether the continued presence of color in the immediate visual environment was necessary for the observation of color-mediated eye movements. Eye movements directed towards the now blank screen were recorded as the sentence unfolded (cf. Spivey & Geng, 2001). In the filler trials, participants looked significantly more at the locations where the targets, rather than the distractors, had been previously presented as the target words acoustically unfolded. In the experimental trials, the locations where the color competitors had previously been presented did not attract increased attention (neither in color nor greyscale trials).

These data demonstrate that language-mediated eye movements are only influenced by color relations between spoken words and visually displayed items if color is present in the immediate visual environment. We conclude that color representations are unlikely to be routinely activated in language processing. Our findings provide strong constraints for embodied theories of cognition which assume that sensory representations influence language processing even in the absence of relevant sensory input. These results fit best with the notion that the main role of sensory representations in language processing is a different one, namely to contextualize language in the immediate environment, connecting language to the here and now.
THE CONSEQUENCES OF MISLOCATED FIXATIONS DURING READING
Laura Wakeford (University of Dundee) & Wayne Murray (University of Dundee)
ljwakeford@dundee.ac.uk

Models of eye movement control during reading differ in terms of whether they assume strictly serial sequential processing of words in text (SAS models, like E-Z Reader, Reiche, 2011) or whether they postulate an attentional gradient that can extend over more than one word, with some parallel lexical processing (GAG models, like SWIFT, Engbert & Kliegl, 2011, and GLENMORE, Reilly & Radach, 2006). It is clear why the duration of fixations might be influenced by lexical properties, like word frequency, of currently-fixated and adjoining words from a GAG perspective. However, SAS models suggest lexical effects should be localised to the currently fixated word (although E-Z Reader contains a mechanism to account for apparent ‘spillover’ effects). It might therefore be concluded that clear evidence for lexical effects before the relevant word is directly fixated – ‘paraparfoveal-on-foveal’ (PoF) effects – provides a critical test of the two classes of model. However, with the suggestion that these could be attributed to ‘mislocated fixations’ (Drieghe et al, 2008), SAS models are also capable of modelling PoF effects. This proposal suggests (a) that a saccade occasionally misses its intended target due to oculomotor error, and (b) the reader, instead of correcting that mislocated fixation, decides to remain at the suboptimal location and nonetheless process the word they are not directly fixating. But while there is much evidence for the occurrence of mislocated fixations, there is rather less for the second postulate. As Engbert and Kliegl (2011) note, such a stay-and-process strategy lies in direct contrast to the error-correcting response that has been proposed to explain the inverted optimal viewing position (IOVP; Vitu et al, 2001) effect, with shorter initial fixation durations near word boundaries attributed to mislocated fixations triggering a fast error-correcting saccade.

The current experiment manipulated the occurrence of mislocated fixations using a novel eye movement contingent text-shift paradigm in which sentence text was moved two characters to the left or to the right during the reader’s saccade from a verb to the following high or low frequency noun. In the no-shift condition, fixations falling on the noun showed a classic IOVP effect, with the duration of accurately-targeted fixations also modulated by word frequency. Following leftward text shifts (causing overshoot) fixation durations were extended somewhat, but followed the IOVP pattern and showed a frequency effect. Critically, in response to rightward shifts (generating undershoots) there was a marked reduction in fixation duration and flattening of the IOVP function, with no evidence of sensitivity to word frequency – all consistent with an immediate error correcting response and inconsistent with a stay-and-process strategy. We conclude, therefore, that the IOVP effect can be related to rapid error correction and there is no evidence in favour of a stay-and-process strategy, at least for fixations that fall short of their intended target.

References
NATIVE AND NON-NATIVE PROCESSING OF ITALIAN SUBJECT PRONOUNS: EVIDENCE FROM EYE-TRACKING

Tihana Kraš (University of Rijeka), Patrick Sturt (University of Edinburgh) & Antonella Sorace (University of Edinburgh).
tk302@cantab.net

Being a null subject language, Italian allows both null (omitted) and overt (expressed) subject pronouns in finite clauses. According to the Position of Antecedent Strategy (PAS) (Carminati, 2002), in intra-sentential contexts the null pronoun is biased towards an antecedent in the subject position, while the overt pronoun is biased towards an antecedent in a non-subject position. The PAS is based on the notion of discourse prominence, defined in syntactic terms: null pronouns are thought to prefer the most prominent antecedents in discourse, antecedents in the subject position being considered as the most prominent ones. Evidence for the PAS has been provided in self-paced reading studies with native speakers featuring pronouns in forward anaphora contexts (Carminati, 2002; Filiaci, 2011).

In an eye-tracking study we investigate whether native speakers (n=28) and proficient non-native speakers whose L1 is English (n=28) process Italian subject pronouns in accordance with the PAS in backward anaphora contexts. Participants read 36 complex bi-clausal sentences with null or overt pronouns in the subordinate clause. The pronoun matched either the subject or the object of the main clause, depending on its gender (Quando lei/lui/Ø è entrata/-o in ufficio dopo pranzo, Adriana ha salutato Roberto con un grande sorriso, ‘When she/he/null entered[fem/masc] the office after lunch, Adriana greeted Roberto with a big smile’). Note that the participal in the initial clause agrees in gender with the subject pronoun (entrata/entrato --- entered[fem/masc]), so that gender matching between the pronoun and the main clause subject can be manipulated regardless of whether the pronoun is overt or null. The backward anaphora context is theoretically interesting because the PAS predicts that the main clause subject (Adriana) will be ignored as a potential antecedent for the overt pronoun due to the hypothesized object preference.

For the native speakers, first pass reading times on the main clause subject plus auxiliary (Adriana ha) show a reliable gender mismatch cost for the null pronoun conditions (mismatch: 320 vs. match: 289 msec), but not for the overt pronoun conditions (mismatch: 310 vs. match: 318 msec), suggesting that the native speakers attempt to assign the null pronoun immediately to the main clause subject, but ignore this position as a potential antecedent for the overt pronoun. Conversely, at the object region (Roberto) a first-pass mismatch cost is found for the overt pronoun (i.e. slow times when the pronoun mismatches with the object; mismatch: 266 vs. match: 243 msec), while the null pronoun conditions do not differ (mismatch: 237 vs. match: 240 msec), suggesting that the native speakers have waited to assign the overt pronoun to this position.

In contrast, the non-native speakers show a very different pattern. First-pass reading times show a mismatch cost for both null and overt pronouns at the subject-plus auxiliary region (null: 385 vs. 362 msec; overt: 379 vs. 355 msec), resulting in a main effect of matching, with no reliable differences at the object region. Therefore, the non-native speakers try to assign both pronoun types immediately to the main clause subject.

To conclude, the native speakers follow the PAS with both null and overt pronouns, while the non-native speakers violate the PAS with overt pronouns, possibly relying on their L1 processing strategies and possibly demonstrating some general properties of L2 processing. Additionally, the native speaker results for the overt pronoun show that a discourse constraint (i.e. the PAS) can lead the processor to ignore the first structurally licit potential antecedent position (i.e. the main clause subject) during active dependency formation.

HOW DO WE COMPUTE PREDICTIONS FOR AN UPCOMING VERB?
Wing-Yee Chow (BCBL; University College London), Cybelle Smith (University of Illinois at Urbana-Champaign), Ellen Lau & Colin Phillips (University of Maryland College Park)
wingyeechow.zoey@gmail.com

Previous research has shown that comprehenders use contextual information to anticipate upcoming input on the fly [1], but less is known about how different sources of information contribute to on-line predictive computations. Recent event-related potential (ERP) evidence indicates that early stages of verb prediction are insensitive to the structural roles of the arguments. The N400, a negative-going ERP component linked to lexical semantic processing, is known to be smaller for more predictable words [1-2]. However, recent studies across different languages have found that reversing pre-verbal arguments in a verb-final clause (e.g., copSUBJ thiefOBJ arrest vs. thiefSUBJ copOBJ arrest), despite clearly affecting the verb's offline predictability (cloze probability), fails to modulate the N400 at the verb [3-5]. A recent study in Mandarin Chinese showed that the N400's sensitivity reemerged only when comprehenders had more time for predictive computations and suggested that the structural roles of pre-verbal arguments have a delayed impact on verb prediction [5]. However, this seems to contrast with a classic finding about the 'verb final' order created by filler-gap dependences in English, where the N400 demonstrated clear sensitivity to the cloze probability of a verb even when it came immediately after the arguments (e.g., John knew which {customer|article} the secretary called…) [6]. Here, to reconcile this apparent discrepancy we explored the possibility that different sources of contextual information impact verb prediction at different time scales. We found that the N400 was sensitive to the cloze probability of a verb when one of the arguments was substituted, but not when the roles of the arguments were simply swapped. We argue that verb prediction immediately and specifically relies on the arguments in a clause, while initially failing to take into consideration the information provided by the structural roles of those arguments.

In Experiment 1 we manipulated the cloze probability of the verb (expected: 26% vs. unexpected: 0%) in object-extracted questions in two ways: in one case cloze was controlled by reversing the same pair of arguments (argument role reversal: a vs. b), in the other case by substituting the extracted argument (argument substitution: c vs. d). All arguments were animate. If comprehenders can immediately exploit all contextual information to anticipate an upcoming verb, then the N400 should be sensitive to the verb’s cloze probability in both cases. Alternatively, if predictive computations cannot immediately exploit argument role information, then N400 insensitivity might be observed. Results (n=24) revealed that only argument substitution elicited an N400 effect. The N400 was completely insensitive to the cloze difference created by the reversal of the argument roles, despite comprehenders' ability to detect the anomaly, as reflected in a P600 effect. Experiment 2 (n=24) showed that argument substitution elicited a significant N400 effect even when the original argument appeared elsewhere in the sentence, indicating that verb predictions are specifically driven by the arguments in the same clause as the verb, rather than by a simple ‘bag-of-words’ mechanism. Taken together, these results show that comprehenders can quickly target the arguments in a clause to predict an upcoming verb, but the arguments’ structural roles have a delayed impact on verb prediction relative to their lexical meaning.

Argument Role Reversal [Exp 1]
(a) The restaurant owner forgot which customer the waitress had served … [26% cloze]
(b) The restaurant owner forgot which waitress the customer had served … [0% cloze]

Argument Substitution [Exp 1 and 2]
(c) The tenant inquired which exterminator the landlord had hired … [26% cloze]
(d) The tenant inquired which neighbor the landlord had hired … [0% cloze]

Chinese relative clauses (RCs) are prenominal. In parsing object-modifying, object-extracted RCs (*John picked up [RC the knife poked] pants*, meaning *John picked up the pants [RC that the knife poked]*), comprehenders need to resolve the ambiguous status of the embN (*knife*) and to attach the headN (*pants*) into the main verb (*picked up*). Existing work using self-paced reading and visual-world eye-tracking has shown that classifiers (e.g., *-tiao*) that mismatch the embN (*knife*) but match the headN (*pants*) could help Chinese comprehenders to avert garden-pathing, though the cueing effect of mismatching classifiers to pre-build RC structure was likely delayed (Wu, Luo, & Zhou, 2014). It is not clear, however, whether the double-inanimate configuration used in that study between the classifier (*-tiao*) and its adjacent noun (*knife*) had induced similarity-based interference (e.g., Gordon et al., 2004), which in turn would have affected comprehenders’ decision to fully establish a clausal boundary. We aim to further probe this issue with Chinese ORCs using ERPs.

**Experiment 1** (N=33; 120 sets of target stimuli) used inanimate emb- and head-Ns (frequency & number of stokes controlled), both being plausible arguments for matrix verbs. We manipulated the semantic congruency of the classifier with i) the (local) embN and ii) the (global) headN (1a-d). We set the (local)match-(global)mismatch condition (1c) as the baseline, because only in this condition can we observe dual effects of mismatching classifiers, specifically, lexical disruption in accessing the embN, and the potential cueing effect for the yet-to-be-seen headN of the target ORC structure. At the embN, local mismatches (1c&d) induced an N400 effect (350-650 ms) compared with local matches (1a&b), suggesting, not surprisingly, lexical disruption effects by local classifier-noun mismatch (Hsu 2006). At the headN, compared with the mismatch-match baseline (1c), both the double-mismatch (1d) and double-match (1a) elicited widely-distributed P200 and left-posterior P600 effects, suggesting difficulties of slotting in the headN.

**Experiment 2** (N=30; 120 stimulus sets) used animate/embN – inanimate/headN configuration, yielding three conditions (2b-d). At the embN, no effects were found across conditions. But at the RC verb, the local match (2b) elicited an N400-like late negativity effect (550-700 ms), suggesting difficulties of repositioning the embN after garden-pathing (due to misparsing the embN as the direct object of the matrix verb). At the headN, the double-mismatch (2d) elicited an N400 effect, suggesting that only when neither the embN nor the headN matched the classifier did comprehenders experience difficulties in slotting in the headN. Such difficulties were absent in the local- or global-match conditions (2b&c), indicating that the target ORC structure was fully constructed prior to the headN, presumably due to the animate embN serving as a good agent perching the subject position of the ORC.

Comparisons of the two experiments suggest that inanimate embN as an implausible agent taking the subject position of the ORC weakened the classifier-noun mismatch cue (Exp. 1), whereas the contrast between inanimate classifiers and animate local nouns successfully led participants not to garden path, and to pre-build ORC structure (Exp. 2). Our results suggest that the cueing strength of mismatching classifier for ORC structure is constrained by the animacy feature shared between the classifier and its adjacent noun.

1. Yanghong | pick up-ASP | (a) that-CL_{knife/pants} [knife] poked [DE] guitar.  
(b) that-CL_{knife/pants} [knife] poked [DE] pants.  
(c) that-CL_{knife/pants} [knife] poked [DE] pants.  
(d) that-CL_{knife/pants} [knife] poked [DE] guitar.

2. Yuhui | talked about|  
(a) Ø  
(b) that-CL_{guest/wine} [guest]brought[DE] wine.  
(c) that-CL_{guest/wine}  
(d) that-CL_{guest/wine} 

Current accounts of sentence comprehension invoke the notion of retrieval interference as a primary determinant of processing difficulty [1-2]. Specifically, similarity between constituents (e.g., NP feature-overlap) has been argued to interfere when people resolve subject-verb or anaphoric dependencies [3-7]. We ask whether similarity-based interference effects arise as a function of multiple NPs in the discourse that overlap in gender and/or number. We take a novel approach by examining interference effects at the second NP rather than downstream after “maintaining” multiple NPs [6-8], using ERPs to establish quantitative and qualitative processing consequences. We used the empty category PRO to introduce two NPs, only the second NP could be PRO controller (e.g., “While [PRO] talking to the waitresses, the man/men/woman/women”). If feature overlap affects processing of the second NP, most interference should occur under gender- and number-matching NPs. Because this interference crosses the subject-object distinction, we predicted that interference would elicit a P600 effect, the effect most reliably associated with syntactic processing difficulties [9].

Methods: During EEG recording, 24 participants read 160 grammatical sentences (40 per condition) in a 2(gender: match, mismatch) x 2(number: match, mismatch) factorial design where the first clause introduced the object-NP and had PRO as subject, and the matrix clause introduced the controller of PRO. Subject and object NPs could overlap in gender and/or number. We fully counterbalanced 160 male/female singular/plural gender-definitional nouns as object NPs, and as critical NP always ‘woman/man/girl/boy’ (or plural form). Sentences were mixed with 156 fillers and presented word by word (300 ms duration, 200 ms blank), followed by intermittent comprehension questions (85% response accuracy).

Results: Across all electrodes, a significant gender by number interaction was observed (500-800 ms window [9]; $F(1,23)=6.02$, $p<.05$), due to a robust P600 effect of number-mismatch in the gender-match conditions ($M=-1.18$, $F(1,23)=8.04$, $p=.01$), that did not occur in the gender-mismatch conditions ($M=-.17$, $F(1,23)=.18$, ns). No distributional effects were observed.

Conclusions: The P600 effect for double-match NPs suggests that interference is driven by similarity contingent upon matching gender and number. Our results testify to the strength of gender-cues during incremental processing, consistent with memory-based accounts of discourse comprehension [2-7]. When features maximally overlap, the subject NP may be momentarily considered as an anaphor for the more distinctive (i.e., first-mentioned and semantically richer) object NP. Alternatively, the P600 may reflect increased discourse complexity stemming from similar NPs [10]. Our results imply a central role for interference during comprehension, even of simple grammatical sentences.

References
[1] Lewis, Vasishth, Van Dyke, 2006
Integrate or Repair? ERP responses to semantic anomalies depend on choice of processing strategy.
Leif Oines & Albert Kim (University of Colorado Boulder)
leif.oines@colorado.edu

Recent ERP studies of sentence processing have found that some types of sentence-embedded semantic anomaly (e.g., “The hearty meal was devouring …”) elicit P600 effects, while many other semantic anomalies (e.g., “The sealed envelope was devouring …”) elicit N400 effects. Theoretical accounts have agreed that such “semantic P600’s” challenge an earlier view that the P600 is elicited specifically by syntactic anomaly, but convergence on a unified, alternative account of the processes underlying P600 and other language-related ERP's remains elusive. Several inter-related but distinct accounts have attributed semantic P600 effects variously to extended combinatory processing triggered by highly anomalous words (Kuperberg, 2007), monitoring in response to conflict between incompatible representations (Kolk & Chwilla, 2007), and the difficulty of semantically integrating highly unexpected words (Brouwer, et al., 2012).

We tested the hypothesis that semantic P600 effects reflect an attempt to structurally repair an anomalous sentence. This response is encouraged by sentences like the first example above, in which “meal” is a plausible Patient of “devouring”, but the sentence is grammatically incompatible with this plausible interpretation (“Semantic Attraction” anomalies). In the sentences without semantic attraction (“No-Attraction” anomalies), the unexpected word is difficult to retrieve from semantic memory, leading to N400 effects.

We asked participants to read these two types of anomalous sentences, and attempted to bias brain responses by manipulating the task that participants performed as they read. All participants received pre-recorded auditory instructions. 22 participants in the “repair condition” were told that some sentences in the experiment were “taken from conversations in which the speaker made an error” and given examples of how such sentences should be “fixed in order for the sentence to make sense”. A separate group of 21 participants (the “integrate condition”) were told that some sentences may have “bizarre meanings” and were instructed to think about what they “could literally mean”. Participants read sentences one word at time from a computer screen, while EEG was recorded from 64 channels at 1000 Hz (Neuroscan SynAmps2).

RESULTS: For participants in the Repair condition, the Semantic Attraction anomalies elicited a robust P600 effect but no N400 effect (see Figure 1 below). No-Attraction anomalies elicited an N400 effect but no P600 effect. For participants in the Integrate condition, neither anomaly type elicited a P600 effect. Instead a left anterior negativity was observed which was larger for the No Attraction anomalies.

Because P600 effects occurred exclusively when participants were instructed to repair anomalous sentences, the results are compatible with a theory that posits structural repair as a key process underlying the semantic P600. LAN-like effects generated by participants instructed to integrate the anomalous sentences could reflect demands on verbal working memory, resulting from attempts to integrate the semantically anomalous representation in which a meal is the Theme of devouring (Kim & Sikos, 2011). Finally, the shift between P600 and LAN effects for the same sentences show that the processes underlying these effects are flexible and modifiable by the comprehenders’ goals.
PROCESSING RECIPROCITY: WHAT INFLUENCES THE STRENGTH OF ‘EACH OTHER’?

Elsi Kaiser (University of Southern California)
emkaiser@usc.edu

Reciprocity is central to human interaction and society (Majid et al.’11), so it is not surprising that languages have different means of expressing reciprocity, such as the English reciprocal ‘each other.’ However, pinning down the meaning of ‘each other’ is challenging (e.g., Dalrymple et al.’98). A sentence like ‘The children pinched each other’ can have a strong reciprocal meaning where each child pinched every other child (see figure), but can also be used in a weaker ‘ring’ configuration where each child pinches someone and is pinched by someone (but not by everyone). Reciprocals also occur in even weaker situations, e.g. ‘chain’ configurations (‘The children followed each other into the room’). (Figs: Majid et al.’11). Conceptually, the strong configuration is more reciprocal than the ring, which is more reciprocal than the chain (where A is only an agent and D is only a patient):

| (1) Strength of reciprocal relation: | [stronger] Strong >> Ring >> Chain | [weaker] |

We investigated what influences whether people construe a stronger or weaker interpretation. Sabato/Winter (2012) hypothesized that lexical meaning, world knowledge and context interact (cf.Dalrymple et al.’98, Kerem et al.’09). We investigated whether the conceptual structure of verbs influences the strength of the reciprocal interpretation that comprehenders construct. Do verbs whose conceptual structure is prototypically asymmetrical (e.g. follow) trigger weaker reciprocal interpretations than neutral verbs (e.g. lick)? Following is prototypically asymmetrical: If X follows Y, typically Y does not follow X (see Dalrymple, Sabato/Winter on a-cyclic graphs). In contrast, licking is neutral: If X licks Y, Y may or may not lick X. If reciprocal interpretation is sensitive to this, asymmetrical verbs (2a) should trigger weaker reciprocal interpretations than neutral verbs (2b).

| (2a) asymmetrical verbs: | The lizards are following/chasing/pursuing each other. |
| (2b) neutral verbs: | The lizards are licking/biting/smelling each other. |

Our secondary, more tentative, aim was to start to explore if the strength of reciprocal interpretations is related to individual differences in empathy. Are higher empathy levels correlated with stronger reciprocal interpretations and a dispreference for configurations where someone is ‘left out’ (e.g. in chain, A and D are not fully engaged)?

EXP-Participants (n=23) arranged sets of small toy animals according to sentences read aloud by a lab assistant facing them (ex.2). Targets used asymmetrical (chase/pursue/follow) or neutral verbs (lick/bite/smell), and consisted of 3 or 4 animals. (We also had trials with 2 animals, as well as various other types and numbers of objects.) Afterwards, people filled out the Interpersonal Reactivity Index (IRI, Davis’80) which measures empathy. Arrangements were videotaped and double-coded afterwards.

RESULTS-We find effects of verb semantics: Neutral verbs result in more (p<.01) ‘strong’ arrangements than asymmetrical verbs (neut=9%, asym=0%); asymmetrical verbs result in more weakly-reciprocal ‘chain’ arrangements (neut=7%, asym=35%, p<.01). ‘Ring’ arrangement rates do not differ significantly. Individual differences: There is a positive correlation (p<.035) between individuals’ IRI scores and preference for ring over chain arrangements. Although this is preliminary (due to small n), it suggests that more empathetic people may show a stronger preference for ring over chain configurations (tend to interpret reciprocals so every animal is both performing and receiving the action; no one is ‘left out’).

To explain strength of reciprocal interpretation, we should consider information about verbs’ conceptual structure, and look more at individual differences in further work: Our preliminary findings suggest interpretation of reciprocals may be linked to cognitive empathy, with more empathetic people tending to opt for stronger (more reciprocal) interpretations.
Language Understanding & Common Sense Reasoning
Joshua K. Hartshorne, Tobias Gerstenberg, & Joshua B. Tenenbaum

Language is frequently ambiguous, with the same sentence having several possible interpretations (The children made delicious snacks). A central challenge for the listener is to determine which of the possible intended meanings the speaker actually meant to convey. One particularly prevalent example is third-person pronouns. In principle, any of the messages in (1) could be encoded in (2):

(1) Al beat Bart at tug-of-war because Al/Bart/Carl/Darrel/Ethelridge/etc. is strong.
(2) Al beat Bart at tug-of-war because he is strong.

However, most people interpret the pronoun in (2) as referring to Al. Intuitively, common sense reasoning is implicated in this inference (cf. Winograd, 1972), though providing an account of common sense reasoning and how it is integrated into language interpretation has proven challenging.

In line with many recent computational models of pragmatics (e.g., Frank & Goodman, 2012; Goodman & Stuhlmueller, 2013) — themselves extensions of earlier theoretical approaches (e.g., Grice, 1989) — we suggest that listeners infer the speaker’s intended meaning using Bayesian inference over an intuitive theory of the speaker’s behavior:

\[
P(\text{message} | \text{utterance}) \propto P(\text{utterance} | \text{message}) P(\text{message})
\]

As a simplification, since pronouns are rarely used to refer to previously unmentioned entities, we assume that \(P(\text{utterance} = (2) | \text{message})\) is negligible for all messages in (1) that do not involve Al and Bart. \(P(\text{message})\) is a function of (at least) what the speaker believes to be true and what is true of the world and the speaker’s experience with it. Note that in de-contextualized sentences like (2), we have no reason to suppose the speaker is more or less interested in conveying any of the meanings in (1), nor do we know what the speaker’s range of experience is. Thus, we need only determine which of the messages in (1) is most likely to be true.

We model tug-of-war competitions as follows. We assume player strength is normally distributed and the stronger player wins (these simplifications can be relaxed but simplify calculation). We model \textit{because} as introducing a counterfactual: \textit{A beat B because A is strong} means that were A not strong, A would not have beaten B. We included 16 variants of (2), manipulating the adjective (\textit{strong, weak}), the verb (\textit{beat, almost beat, lost to, almost lost to}), and the connective (\textit{because, although}). \textit{Strong} and \textit{weak} are modeled with prototype semantics. We model \textit{although q} as meaning \textit{because of q, p was unlikely}. The model’s and humans’ interpretations of the pronouns correlated well (\(r=.92, p<.0001\)).

In order to ensure that our results were not specific to the tug-of-war scenario, we tested an additional 40 pairs of sentences involving explanations (3a) and 40 pairs involving results (3b):

(3) a. Al frightened Bart because he is reckless/timid.
   b. Because Al frightened Bart, he got in trouble/ran away.

In each pair, the sentences differ in terms of the most likely referent. Because we do not have a full model of the world from which to derive the probability that Al or Bart being reckless would cause Al to frighten Bart or the probability that Al frightening Bart would cause Al or Bart to get in trouble (etc.), we asked a separate set of participants to rate those probabilities, which were then used in the language model. The correlations between the model’s and humans’ pronoun interpretations was high for both sentences like (3a) (\(r=.88, p<.001\)) and (3b) (\(r=.73, p<.001\)).

We discuss this model and these findings in the context of recent work on the role of discourse structure, syntactic structure, and verb biases in pronoun interpretation (Hartshorne & Snedeker, 2013; Kehler & Rohde, 2013; Sagi & Rips, in press).
INCREMENTAL INTERPRETATION AND REFERENCE TO CONTRAST SETS
Craig Chambers & Daphna Heller (University of Toronto)
craig.chambers@utoronto.ca

Theoretical and experimental work has often appealed to the notion of a contrast set to explain aspects of semantic and pragmatic interpretation. For example, when listeners interpret the expression the tall glass, they not only identify a referent but also identify another member of the set with which the referent contrasts (Sedivy et al., 1999). Other phenomena involving contrast sets include focus operators such as only, quantifiers such as few of, and prosodic focus (e.g., Ni et al., 1996; Moxey & Sanford, 1987; Rooth, 1992). Unlike these cases, where the existence of a contrast set is implicitly conveyed, expressions such as the other X refer directly to a member of a contrast set, with the adjective not indicating any particular properties of the referent except its membership in the contrast set. The current study aims to exploit these properties of the expression the other X in order to examine the mechanisms underlying the computation of contrast sets.

Theoretical studies have suggested that other functions in an anaphorically straightforward way, such that the interpretation of this adjective in a linguistic description is semantically tied to the nominal category evoked in an earlier description (Umbach, 2002). On this account, hearing other should lead listeners to expect a referent that belongs to the nominal category mentioned previously. However, listeners might also compute a broader set of relationships that could plausibly reflect other types of links among referential entities. In the current experiment, we test these possibilities by examining real-time interpretation in a visual world experiment where on each trial listeners followed a sequence of instructions to click on depicted objects within a four-image display.

On critical trials, auditory materials had the form: "In the top right corner there is an [N1]. Click on the [N1]. Now click on the other [N2]". Three conditions were used. In the same-category condition, the same noun (e.g., bat) was used in both instructions, and the display included two instances of the same basic-level category (e.g., two instances of the animal bat), along with two unrelated images; this corresponds to the straightforward anaphoric situation described earlier. In the superordinate condition, [N1] remained the same but [N2] was a superordinate term (e.g., animal), referring to another instance of the same higher-level category (e.g., a cow); this enabled a test of whether listeners go beyond the mentioned noun and use conceptual relationships to rapidly define a contrast set. In the homophone condition, [N1] and [N2] were the same, as in the same-category condition, but the display contained a homophonic image (e.g., a baseball bat); this provided a test of the role played by the mentioned noun, asking whether contrast sets might be defined on the basis of linguistic form even when the conceptual relationship is entirely ad hoc.

To assess the on-line computation of contrast sets, we examined eye movements during the processing of other, before any information from the noun was available. In the same-category condition, listeners anticipated the intended referent, reflecting the immediate computation of a contrast set based on the previously-mentioned noun. Interestingly, listeners also anticipated the referent (albeit to a weaker extent) in the superordinate condition, suggesting that they were computing a contrast set based on conceptual relationships as soon as other was heard. Strikingly, listeners also anticipated the referent in the homophone condition, where there is no conceptual relationship between the two entities, and the only link involves phonological overlap in their nouns. Together, the results suggest that listeners can readily compute a contrast set based on conceptual or form-based relationships among entities. These results underscore the highly flexible and inferential nature of referential processing during real-time comprehension.
COGNITIVE LOAD AFFECTS THEORY OF MIND-USE IN THE DIRECTOR TASK

Jessica Wang (University of Birmingham), Steven Frisson (University of Birmingham), & Ian Apperly (University of Birmingham)
j.j.wang@bham.ac.uk

In the director task (Keysar et al., 2000, 2003), objects are placed in a grid with a “director” standing on the other side of the grid giving instructions to move objects. Some objects are blocked from the director’s view and only visible to the participant. In order to follow the director’s instructions correctly, participants have to take into account the director’s perspective, i.e. they need to be able to think about others’ perspectives, using their Theory of Mind. Several studies (see Apperly et al., 2010, for an overview) have shown that adults frequently make egocentric errors in such a task. Moreover, eye movements show that participants often fixate a privileged object earlier and for longer than when the same object is in common ground (but see Brown-Schmidt, 2009; Hanna et al., 2003).

We report a number of director task studies in which we measured both accuracy and speed of response and eye movements during instruction-comprehension and referent-selection, examining to what extent different task demands impacted upon egocentric behaviour. We systematically varied three possible sources, corresponding to different demands one faces in everyday communicative situations.

First, we found that the type of instructions can greatly affect the amount of egocentric errors: when participants had to infer the context in which their knowledge about perspective-difference had to be applied, they committed egocentric errors on about 40% of the critical trials. When participants received extensive instructions including an example of how referent-selection may differ according their own perspective and the director’s perspective, the error rate dropped to 2%. However, even with this low number of errors the experimental condition engendered more errors and longer reaction times than the control condition, where it was not necessary to consider perspective information.

Second, the number of objects in common and privileged ground only affected overall reaction times but not accuracy or degree of egocentrism.

Third, when participants received more complex syntactic instructions (“nudge the small ball one slot up” vs. “nudge the small ball”) they were slower to make a deciding fixation on the target in the experimental condition than the control condition. This was likely due to the demand on syntactic processing competing with the resources required to overcome egocentric tendencies.

To conclude, the cognitive load associated with language comprehension can affect the amount of egocentrism observed. The demand on both pragmatic and syntactic levels had clear effects on participants’ ability to select the correct referents. In contrast, the visual aspect of the task i.e., the magnitude of the common versus privileged ground had little direct effect on overcoming egocentrism. Such systematic analysis highlights the factors one ought to consider during communication.


Referring expressions like personal pronouns can be used to refer back to entities previously mentioned in the discourse. The more potential antecedents a referring expression has, the more difficult it is to resolve (Ariel, 1990). This was shown in Dutch for potentially ambiguous overt pronouns compared to less ambiguous full noun phrases using pupil dilation as a measure of cognitive effort (authors, submitted).

In contrast to speakers of Dutch, speakers of the pro-drop language Italian can use either an overt pronoun (e.g., lui ‘he’) or a null pronoun (Ø) as a subject anaphor. In Italian, null pronouns generally refer to the discourse topic, whereas overt pronouns refer to a non-topic referent (Carminati, 2002). However, this is merely a preference and interpretations of null as well as overt pronouns can vary. According to Carminati (2002), the interpretation preference of overt pronouns is weaker than that of null pronouns. If there is such a difference in interpretation preference, can we see this difference in the online processing of anaphoric subjects too? We predict that the amount of cognitive effort needed to resolve an anaphor is related to its ambiguity: the weaker the interpretation preference of an anaphoric expression, the more effortful its resolution. Thus, we predict the more ambiguous overt pronouns to elicit a larger pupillary response than null pronouns.

To test this prediction, 40 Italian adults heard short stories containing one of three different subject anaphora: A full NP such as the hedgehog as an unambiguous baseline condition, a null subject pronoun, and the overt subject pronouns lui (‘he’) and lei (‘she’). We determined participants’ final interpretations and measured pupil dilation during processing.

Analyses of the offline interpretations showed that null pronouns were mostly interpreted as referring to the discourse topic, whereas overt pronouns generally referred to a non-topic referent (p < 0.001). Moreover, the interpretation preference for overt pronouns was weaker than for null pronouns (resp. 60% and 85%). The proportional pupil dilation over time, aligned to the offset of the subject, is plotted in Figure 1. An analysis of the pupil dilation data over time was performed with Generalized Additive Models (GAMs; Wood, 2006). The analysis revealed that null and overt pronouns evoked more pupil dilation than full NPs (both p < 0.001), and that overt pronouns evoked more pupil dilation than null pronouns (p < 0.001).

Figure 1: The proportional pupil dilation over time, aligned to the offset of the subject anaphor. A higher proportional pupil dilation reflects the investment of more cognitive effort.

From these results we conclude that ambiguity resolution in Italian subject anaphora is reflected by pupilary measures of cognitive effort. Moreover, the pupillary data support the prediction that the amount of cognitive effort needed to resolve an anaphor is related to its ambiguity: overt pronouns, which showed a weak interpretation preference, are more effortful to resolve than null subject pronouns with a stronger interpretation preference.
BIDIRECTIONAL SYNTACTIC PRIMING IN CONVERSATION: I AM PRIMED BY YOU IF YOU ARE PRIMED BY ME

Lotte Schoot, Peter Hagoort & Katrien Segaert (Max Planck Institute for Psycholinguistics)
Lotte.Schoot@mpi.nl

We investigate whether the magnitude of syntactic priming effects, like other aspects of linguistic behavior, is mimicked between speakers in a conversation. Speakers adapt syntactic structures to their interlocutor (i.e. syntactic priming), but what is more, speakers adapt the magnitude of syntactic priming effects to interlocutor's magnitude (Schoot et al. 2014). We expect to replicate and extend this finding with the present results.

60 participants were divided into 30 pairs who performed the experiment together. They were asked to describe photographs to each other, depicting two persons performing a transitive action (e.g. a man hugging a woman). Descriptions were actives or passives (no free choice, see Menenti et al., 2011). Priming effects were measured by comparing speech onset latencies for sentences with repeated syntax (two consecutive actives or passives) relative to novel syntax (active follows passive or vice versa). Before participants performed this communicative task, we ran a non-communicative pre-test for each participant, to measure their individual priming effect without influence of the partner's priming effect.

The results show that priming magnitude is determined by your partner's priming effect; the more your partner is primed by you, the more you are primed by your partner \((r = 0.572, p < 0.002, \text{Figure A})\). Furthermore, the difference between paired speakers' individual syntactic priming effects (as measured in the pre-test) predicted how much speakers adapt their syntactic priming effects when they are communicating with their partner in the communicative experiment \((\beta = -0.632, p < 0.001, \text{Figure B})\). If in the pre-test, your partner is primed more/less by you than you are by her, you will increase/decrease your own priming magnitude towards your partner's priming magnitude in the communicative context.

Syntactic priming effects in conversation are said to result from speakers aligning their syntactic representations by mimicking sentence structure (Pickering & Garrod, 2004; Jaeger & Snider, 2013). Here we show that on top of that, the magnitude of priming effects is also mimicked between interlocutors. Although measuring syntactic priming effects in speech onset latencies is now proven to be a valid method (Corley & Scheepers 2002; Segaert et al. 2010; 2014; Smith & Wheeldon 2000; Wheeldon & Smith 2003), we are currently running an experiment to investigate whether we can replicate this effect when looking at syntactic priming magnitude in structure choices.

Figure A. The syntactic priming effect (ms) of one speaker in a pair is significantly influenced by the priming magnitude of this speaker's communicative partner. The more your partner is primed by you, the more you are primed by your partner. Speaker A is always the participant with the smallest priming effect. B. The more different a speaker's priming effect (ms) is from her partner's in the non-communicative pre-test, the more this speaker's priming effect changes when she is communicating with this partner. If your partner's priming effect is bigger than your own in the pre-test (negative x-value), your priming effect increases in the communicative experiment (positive y-value) and vice versa. Speakers thus (unconsciously) adapt their own priming effects towards their partner's priming magnitude in conversation.
THE EFFECT OF MOTIVATION AND WORKING MEMORY LOAD ON PERSPECTIVE-USE: EVIDENCE FROM EYE-TRACKING.

James Cane (University of Kent), Heather Ferguson (University of Kent) & Ian Apperly (University of Birmingham)
j.e.cane@kent.ac.uk

Language comprehension frequently requires the listener to adopt a speaker’s perspective in order to narrow down an intended target referent. For example, to understand a speaker who asks for “the ball” when more than one ball is available, one must infer that speaker’s knowledge of, or access to, the available objects. Over the last decade, a number of psycholinguistics researchers have used the visual-world paradigm to investigate how such inferences evolve over time during communication (see Huettig, Rommers, & Meyer, 2011). Recently, research has demonstrated a link between the ability to use perspective during language processing and working memory (WM; e.g. Lin, Keysar, & Epley, 2010). Across two eye-tracking experiments we examine whether motivation modulates the effects of WM load on the time-course of perspective-taking in a referential communication task.

In the task participants moved target objects (e.g. a glass with a straw in) around a grid based on instructions from an avatar (e.g. “Move the glass with the straw in left”). In a ‘listener privileged’ condition, an alternative target object (e.g. a glass with a lemon in) could only be seen by the participant. In a ‘speaker privileged’ condition, there was an indication of a potential target object only available to the speaker. In a control condition, both target and alternative target objects were available to both participant and speaker. Working memory load was manipulated within each condition. In Experiment 2 (N=31), but not Experiment 1 (N=36), participants received financial reward for quick and accurate responses.

In Experiment 1 - where there was no reward or time-pressure - listeners did not use perspective cues to disambiguate the target object from the distractor under either the high or low WM load. In Experiment 2 - where there was a reward for speed and accuracy - listeners used perspective cues to disambiguate the target object from the distractor object from the earliest moments of processing, but only under low load. Under high load, responses were similar to that of the control condition, where both objects were in common ground. Furthermore, attempts to initiate perspective relevant responses under high load led to reduced recall on the concurrent WM task – indicating that inferring perspective from a speaker’s utterance was drawing on limited cognitive resources. Across both experiments listeners’ fixations were not influenced by the speaker’s potential knowledge about an alternative target.

These results show that, when there is no explicit motivation to use perspective, holding privileged knowledge about objects interferes with our ability to take another’s perspective to narrow down reference. However, when there is sufficient motivation to use perspective, the ability to ignore distracting objects is ameliorated, but only under low cognitive load. This indicates that perspective cues can guide referential interpretation of language, but only when there is sufficient motivation and sufficient cognitive resources to do so.

References:

LEARNING AND COMMUNICATION JOINTLY ALIGN PAIRS’ CATEGORY SYSTEMS

Catriona Silvey, Simon Kirby, Kenny Smith (University of Edinburgh)
C.A.Silvey@sms.ed.ac.uk

The interactive-alignment model predicts that communication leads to alignment of individuals’ representations (Garrod & Pickering, 2009). Previous work tested this hypothesis in an experimental paradigm where solitary individuals and communicating pairs divided a continuous space of morphed images into labelled categories (Silvey, Kirby, & Smith, 2013). Surprisingly, the category systems of communicating pairs were significantly less aligned than the systems of matched pairs of individuals who did not interact. The difficulty of coordinating on conventions led communicating pairs to focus agreement on particular salient exemplars, leaving their overall category systems relatively unaligned.

In the real world, however, communicators do not coordinate on conventions from scratch: they acquire shared conventions from a language passed down over cultural transmission. This language provides common ground, which may be a necessary pre-condition for interactive alignment. We ran an experiment where pairs first learned a category system before communicating. Participants were assigned to chains of 5 pairs. The first pair in each chain learned a random labelled category system. This pair then modified their category systems by playing a communication game, as in Silvey et al. (2013). The resulting category systems were passed on to be learned by both members of the next pair in the chain. Over generations, the alignment of successive pairs’ category systems gradually increased, until generation 5 pairs achieved the same level of alignment as matched pairs of individuals from experiment 1. This gradual increase in alignment came about via both a) learning biases amplified by cultural transmission and b) interactive alignment. Learning pressures caused category systems to simplify, making them easier to learn in later generations and therefore increasing the baseline alignment for these pairs. Importantly, however, the cumulative increase in alignment was not only due to learning. Consistently over generations, pairs’ category systems were more aligned after communication than after learning alone. Thus, learning and communication appear to have differently beneficial effects on alignment. Participants’ errors in learning the category systems simplify them in accordance with general human learning biases; these more learnable systems provide common ground, which communication then works to align further within pairs.

We also measured the average alignment of category systems between pairs by comparing participants from different chains who did not communicate together. Between-pair alignment also increased over generations, suggesting some level of convergence towards universal category structures. However, alignment was consistently higher within pairs than between pairs, showing that the category system on which a given pair will align is partly contingent on that pair’s particular interaction history. The results illuminate the role of learning and cultural transmission in providing the necessary common ground for interactive alignment, and in producing the mixture of universal and language-specific category structures we see in natural languages.

References


Investigation of Chinese-English learners’ acquisition of tense and aspect in English: self-paced reading and eye-movement study

Panpan Yao\textsuperscript{a} (Queen Mary, University of London), Baoguo Chen\textsuperscript{b} (Beijing Normal University)
\textsuperscript{a} p.yao@qmul.ac.uk \textsuperscript{b} Chenbg@bnu.edu.cn

L2 acquisition of tense and aspect has been extensively investigated, but most research relies on corpus studies and off-line tasks like acceptability judgments, providing a picture of L2 learners’ grasp of explicit knowledge ([1], [2], [3]). What is less clear is whether L2 learners can use grammatical knowledge unconsciously and automatically in real-time comprehension tasks. We present experiments employing self-paced reading and eye-tracking to investigate how late Mandarin Chinese-English (C-E) learners process tense and aspect in English. Progressive aspect is grammaticalized in MC, while the past/present tense distinction is not (following [4]'s criteria). By comparing C-E learners’ on-line processing of English verb forms, we look into the potential role of L1 in the acquisition of tense and aspect in L2.

[5] used self-paced reading to investigate whether advanced French and German learners of English are sensitive to tense/aspect mismatches between a fronted temporal adverbial and the verb form in on-line comprehension. They find that, despite performing similarly in explicit tasks, only those who’s L1 has grammaticalized aspect (French) were sensitive to the tense/aspect violations on-line.

Our Exp1 uses a similar self-paced reading design (also with fronted adverbials) to investigate whether C-E learners can detect errors in progressive, simple past and present third person singular forms in English when they were asked to comprehend sentences. Materials were as in Table1. Both high and low proficiency C-E learners rapidly detect errors in progressive verb forms in English, but only high proficiency C-E learners detect errors in the simple past. This result differs from [5]’s in that C-E learners were sensitive to the past tense verb form violations on-line even past tense is not grammaticalized in MC.

Exp2 used eye-tracking to provide a more fine-grained measure of processing. We found that both high and low proficiency C-E learners rapidly detected errors in the progressive, but only high proficiency C-E learners detected errors in the simple past and present third person singular forms, confirming the results of Exp1.

Our results support the Full Transfer/Full Access Model ([6]) in which the L1 has an influence on L2 learners’ on-line processing of tense and aspect: a grammatical feature active in L1 will be easier to acquire and process in L2. However, L2 learners can ultimately acquire tense and aspect with high proficiency levels, even when these features are not grammaticalized in L1.

<table>
<thead>
<tr>
<th>Table 1: Sample Materials</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Material type</strong></td>
</tr>
<tr>
<td>V-ed</td>
</tr>
<tr>
<td>V-ing</td>
</tr>
<tr>
<td>V-s</td>
</tr>
</tbody>
</table>

WORDS ARE LESS SALIENT IN PHRASES: EVIDENCE FROM THE FREQUENCY EFFECT IN RECOGNITION MEMORY

Cassandra L. Jacobs, Gary S. Dell, Aaron S. Benjamin (University of Illinois) cljacob2@illinois.edu

Normally, high frequency words are easier to produce and comprehend than low frequency words. However, a word frequency paradox arises in tests of recognition memory, where low frequency words are easier to recognize as having been in a study list than high frequency words [2]. This may be because common words have more episodic representations, making it difficult to recover any particular one when memory is tested [4]. Jacobs et al. (2013) examined whether memory for low frequency phrases behaves similarly. Participants studied adjective-noun phrases varying in their frequency (e.g. “alcoholic beverages”, “flaming bounds”). They found that while people can assess how common a phrase is, their accuracy for picking out phrases did not depend on phrase frequency. Instead, there was a hint that accuracy depended on how common the nouns within them were. Specifically, if a noun within a phrase was infrequent, the phrase was a bit more likely to be accurately classified as old or new. They proposed a model where phrases are stored and remembered compositionally. Because the meaning of phrases is computed from its words and stored compositionally, word frequency trumps phrase frequency in affecting recognition memory.

We replicated and expanded on these results by looking at whether participants would remember adjective-noun phrases best when they contained low frequency nouns. We first selected high and low frequency nouns from [1], e.g. “wizard” or “tree”. Using these as single-word stimuli in Experiment 1, we replicated the low frequency advantage in recognition memory, where studied low frequency words are more likely to be recognized and unstudied ones less likely to be falsely recognized. In short, “wizard” was better remembered than “tree”. Then, for Experiment 2, we placed the nouns in adjective-noun phrases not varying in phrasal or adjective frequency (e.g. “handsome wizard” or “premature tree”). The participants studied and were tested on phrases. We found that participants were able to make more accurate recognition judgments when the phrases contained low frequency nouns, but the low frequency advantage was significantly attenuated from Experiment 1 (see Table). In Experiment 3, we tested whether the attenuated effect in Experiment 2 was due to processing during study, or processing of the phrase at test. Participants studied the phrases from Experiment 2 but were tested only on their memory for the nouns, as in Experiment 1. Participants' accuracy on individual nouns varied more strongly as a function of noun frequency than in Experiment 2. Even though they studied phrases, their ability to recognize the nouns out of context was based on properties of the nouns. We attribute the smaller effect in Experiments 2 and 3 to initially encoding phrases at study.

Overall, an individual word may contribute less to long-term memory when it is presented in a phrase at either study or at test. Phrases by themselves contain more information than individual words, because meaning can be composed from them and because they inherit the meanings of their words. For example, “handsome wizard” could bring up the meanings of handsome, wizard, and the image of Harry Potter. Our results also support the model of Jacobs et al. (2013) by showing that individual word frequencies can by themselves influence our memory for phrases when that is the primary source of information about a phrase, but that other kinds of information are available to memory about phrases that are not there for nouns.

References

<table>
<thead>
<tr>
<th>Size of the low frequency word advantage in recognition memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

β represents decrease in accuracy in log odds per unit of log₂ frequency
The English past tense has had a long and troubled history as a major psycholinguistic battleground between opposing views of mechanisms of language processing and cognitive structure (McClelland & Patterson, 2002; Pinker & Ullman, 2002). In the current study we seek to open a new dimension in this debate by focusing on the role of memory systems in learning novel morphological forms. We examine the predictions of a Complementary Learning Systems approach (McClelland et al. 1994) which suggests that the similarity between newly acquired information and existing knowledge crucially determines the relative involvement of the hippocampal vs. neocortical memory systems in acquiring new knowledge. In particular, high levels of similarity between the newly learned and existing memories should allow swift integration in the long-term neocortical stores, whereas low levels of similarity will require greater hippocampal involvement and thus be subject to a stronger influence of memory consolidation.

We designed a set of novel verbs that were phonologically similar to existing English verbs: half of the items were similar to neighborhoods of existing verbs with predominantly regular past tense forms (e.g. PLARE; cf. share-shared, stare-stared), and the other half to neighborhoods of existing verbs with predominantly irregular past tenses (e.g. FLEEP; cf. sleep-slept, keep-kept). Participants (native English speakers) were first trained on the present tense forms of the novel verbs (PLARE, FLEEP) using word repetition in a sentential context, and sentence completion. A week later they were trained on the past-tense forms of the novel verbs: these were either consistent (e.g. plared, flept) or inconsistent (e.g. plore, fleeped) with the existing neighborhoods. Participants were trained on the past-tense forms in the morning or in the evening and were tested 12 hours later using past tense generation, recall and recognition tasks.

When tested immediately after learning, the generation of the novel past tense forms was influenced by both the regularity of the existing neighborhood and whether the given past tense was consistent with the properties of the neighborhood, as well as the regularity x consistency interaction. Crucially, there was a significant difference between sleep and wake in the performance at the delayed test (figure on the left): while for the participants who slept between training and test the memory for both the novel forms consistent and inconsistent with the existing predominant pattern (regular) was preserved, for the participants tested after a wake delay only the memory for the consistent forms was significantly impaired (plared, fleeped), whereas the memory for the inconsistent forms was preserved (plore, flept).

These findings provide clear evidence in support of the hypothesis that new linguistic forms consistent with the existing knowledge can be rapidly acquired via the support from the long-term neocortical knowledge (McClelland, 2013). New forms inconsistent with the existing knowledge are subject to a stronger influence of memory consolidation due to an increased involvement of the hippocampal system. More broadly, these findings suggest a role for domain-general memory mechanisms in learning novel morphological forms.
The effect of literacy on phonological processing has been described as analogous to a virus that "infects all speech processing" (Frith, 1998). Behavioural data has shown that exposure to literacy training in alphabetic languages coincides with both qualitative and quantitative improvements in the awareness of the phonological structure of words. Recent brain imaging data indicates that such behaviour may result from a restructuring of phonological processing networks as a consequence of learning to read an alphabetic language. Harm & Seidenberg (1999) present a computational model of reading that provides an explicit description of how such phonological restructuring may occur. They showed that connectionist networks trained to map between English orthographic and phonological representations displayed more componential phonological processing than networks trained only to stably represent the phonological forms of English words. This model and processing level models such as Psycholinguistic Grain Size Theory (Ziegler & Goswami, 2005) predict that systems trained on alphabetic languages will develop more componential processing due to the systematic relationships that exist between the letters of the language and corresponding speech sounds.

We trained a model similar to that described in Harm & Seidenberg (1999) to capture explicitly how the effects of literacy training on phonological processing may differ as a consequence of orthographic transparency. We developed two models, a transparent literate model and a non-transparent literate model, both of which were trained on a corpus of 6,188 English words with orthographic and phonological representations. For the transparent model the mapping between orthographic and phonological representations corresponded to the English mappings, however in the corpus used to train the non-transparent literate model orthographic representations were randomly assigned to phonological representations (e.g. Transparent: cake -> keɪk, fort -> fɔːt; Non-transparent: cake -> fɔːt, fort -> keɪk). Thus the non-transparent model had to learn the relationship between the whole word and its pronunciation without recourse to regularities at a finer grain-size. This arrangement ensured that the two models were controlled in terms of the set of inputs and outputs but differed only in the extent to which the mapping was transparent.

We observed that networks trained on a transparent orthography were better at restoring phonetic features and phonemes, reflecting componential phonological processing. However, networks trained on a non-transparent orthography were more likely to restore corrupted phonological segments with legal, coarser linguistic units (e.g. onset, coda). Our results connect with the growing body of work that describe differences in phonological processing between Chinese (a morphosyllabic language) and English literate populations and provides an explicit description of, and predictions for, how differences in orthographic transparency can determine varying strains and symptoms of the 'literacy virus'.

Alastair Smith (Max Planck Institute for Psycholinguistics), Padraic Monaghan (Lancaster University) & Falk Huettig (Max Planck Institute for Psycholinguistics).

alastair.smith@mpi.nl
Recent findings indicate processing difficulties for German masculines both in aphasia (Seyboth, Blanken, Ehmann, Schwarz, Bormann, 2011) and with healthy L1 adults (Opitz, Regel, Müller & Friederici, 2013). In three experiments we explored the L2 incidental acquisition of grammatical gender of novel German nouns (NNs) in order to find out whether there are acquisition differences between the three genders in L2 as well.

In a pilot study, advanced learners of German with Slavonic L1s read short texts, each with three occurrences of a NN presented within unambiguously gender marked NPs. The 20 NNs replaced existing low frequency German nouns and were assigned either the masculine, feminine, or neuter gender. After each text, several additional sentences were read in a self-paced manner (SPR). One of them contained the NN within a gender marked phrase which was either gender-congruent (congruent condition) or gender-incongruent (incongruent condition) with the gender assigned to the NN in the text. The results revealed significant Plausibility x Gender interaction: Feminine NNs were significantly slower (38ms) in the incongruent condition and the same tendency was observed for neuter nouns indicating that participants acquired these two genders and perceived their occurrence within gender incongruent NPs as violations. On the other hand, masculine NNs exhibited a numerical tendency of 18 ms in the opposite direction.

In Experiment 2 we explored the gender acquisition with carefully controlled materials without manipulating additional factors as it was the case in the pilot experiment. Based on a rating study we selected 24 NNs (eight of each gender) such that each of the NNs was rated equally probable for two genders. In the SPR experiment, a given NN was presented with one of the two genders in the text and with the other gender in the incongruent condition for half of the participants. For the other half the combination with the two genders in the text and in the incongruent sentence was reversed. This balanced design guaranteed that there was no bias towards a particular gender based on the phonological form of a NN. The pattern of results was the same as in the pilot study: Feminine and neuter NNs were significantly slower when presented within gender incongruent NP in the SPR sentence, while masculine NNs were faster.

The results of the SPR experiment were confirmed also by a grammaticality judgement experiment that followed Experiment 2. Participants had to decide whether NPs consisting of a definite article and a noun were grammatically correct. Each NN was presented twice in a balanced design, once with a correct and once with an incorrect gender. Whereas participants were significantly faster when accepting correct NPs involving familiar existing words than refusing such incorrect NPs, this standard pattern could be replicated only for feminine and neuter NNs. For masculines, the RTs of yes- and no-answers did not differ. The acquisition differences between the three genders shall be discussed in the context of similar earlier findings (Mills, 1986) as well as in the context of the processing differences (Opitz, et al.2013). The latter ERP study addresses the topic from the perspective of underspecification-based theories and mirrors our results in that no LAN-effects (marking morphosyntactic violations) for masculine NPs with gender mismatch were observed, in contrast to the two other genders.

PHONOTACTIC PROBABILITY IN ADULT L1 AND L2 WORD LEARNING

Amit Kirschenbaum, Denisa Bordag, Maria Rogahn, & Erwin Tschirner
(University of Leipzig)
denisav@uni-leipzig.de

In general, phonologically regular words are more easily stored in memory than irregular ones (e.g. Ellis & Beaton, 1993). On the other hand, readers must direct their attention to the new word in order to acquire it incidentally, and phonotactic atypicality can contribute to this noticing process (Schmidt, 2012). In four experiments we explored the role of phonotactic probability in incidental and intentional acquisition of new words in L1 and L2 German.

Twenty phonotactically typical and 20 atypical (but phonotactically possible) pseudowords were selected in a pre-test and pairwise matched. The pair members differed in their phonotactic typicality quantified by bigram and trigram transition probabilities, and L1 and L2 ratings, but otherwise had the same relationship to existing German words (neighbourhood density and orthographic-Levenshtein distance 20).

Participants first read short texts, each with three occurrences of a novel word (the pseudoword) substituting a low frequency noun (“Speue/Gnia” – i.e. ‘aviary’). Half of the novel words were typical (e.g. “Speue”), the others atypical. For the second half of the participants, the other pair members were inserted in the texts (e.g. “Gnia”). After each text, several additional sentences were read in a self-paced manner, one of which contained the novel word preceded by either a semantically plausible (e.g. renoviert – ‘renovated’) or implausible adjective (e.g. schnell – ‘fast’). In the intentional version of the experiment, participants studied the novel words from definitions.

The L1 results on intentional learning showed slower reading times of the novel words (and spillover regions) in the implausible condition irrespective of their phonotactic probability. In the incidental learning, however, the L1 participants reacted earlier to the semantic violation in the implausible condition when the novel words were phonotactically atypical, indicating that their meaning representations were available earlier and probably more strongly represented. This supports the hypothesis that phonotactic salience initiates noticing in incidental learning. On the other hand, in L2 intentional learning, the implausibility effect was larger for the phonotactically typical than atypical novel words and of the same size for the two groups of novel words in the incidental learning.

The results of the following Repetition Priming Experiment revealed a recognition advantage (larger repetition priming) for the atypical novel words across all tasks, supporting recent evidence from child language acquisition indicating that rare sound sequences facilitate the cognitive process of triggering (Storkel & Lee, 2011). The results of the SPR experiment, however, show that incidental and intentional learning in L1 and L2 are affected by various determinants of word learning to a different degree. The results will be discussed with respect to the role of the phonological loop in vocabulary acquisition and its interaction with the phonological information stored in long term memory as determined by the size of the L1 and L2 lexicons (Munson, 2001). The role of attention and sublexical and lexical processing (Vitevitch & Luce, 1998) will be addressed as well.

In two experiments we explored the processing of morphologically complex words in Czech, a Slavonic language with rich inflectional system with frequent prefixes and suffixes. The focus was on the processing of inflectional and derivational forms (Exp. 1) and on the impact of morphological transparency/opacity on the processing of prefixed verbs (Exp. 2).

In Experiment 1, we employed morphological repetition priming (identity, inflectional, derivational) to investigate whether Czech derived and inflected word forms are stored in the mental lexicon in the same or different manner (cf. the Split Morphology Hypothesis, Anderson, 1988). The target was always either a verb in 1.Sg.Pre.Ind., e.g. PLAVU (I swim) or a noun in Nom. Sg. (HMAT – a touch). The prime was either identical, or inflectionally (PLAVEŠ – 2.Sg.Pre.Ind.; HMATEM – Instrumental Sg.), or derivationally (PLAVEC – a swimmer; HMATÁŠ – you touch) related. Participants performed lexical decision on the targets. Fillers and pseudowords (partly constructed analogically to the experimental items, e.g. PIRU, PIREŠ, PIREC) were also included. The results revealed significantly smaller derivational priming effects compared to inflectional and identity priming (which were statistically the same). There was no difference in the no-responses to the pseudoword groups that were constructed analogically to the experimental items. The pattern of results indicates distinct representations of inflection and derivation in the mental lexicon, with inflected forms being fully decomposed and derived forms either not or only partially.

The main goal of Experiment 2 was to explore the storage of the derivational forms in more detail by addressing the role of semantic transparency in the recognition of prefixed verb forms. In addition, we decided to test the psycholinguistic reality of two proposed functions of verbal prefixes: the "purely aspectual function" (the prefix turns an imperfective verb into a perfective one without change of lexical meaning and can be seen as inflective, e.g. HŘEŠIT (imperfective) - ZHŘEŠIT (perfective), to sin) and derivational prefixes that also change the verb meaning (e.g. KRESLIT (imperfective), to draw – ZKRESLIT (perfective), to distort). The existence of such a distinction is a subject of ongoing debate in the Czech linguistics.

In order to explore these topics, we compared the response latencies (lexicality decision) to three groups of verbs that were controlled for length and frequency: A) verbs with presumably inflectional aspectual prefix, B) semantically transparent verbs with derivational prefix, C) semantically opaque verbs with derivational prefix. The prefixes were in all groups z- and u- (i.e. they were homonymous with respect to the three functions).

No difference was found between the groups A and B. Thus, we found no support for the mental existence of two distinct functions of verbal prefixes which would correspond to the difference between inflection and derivation. However, slower RTs for group C compared to A and B confirms findings in other languages that semantic transparency affects lexical storage and access of morphologically complex word forms and indicates that semantically transparent derived forms are at least partly decomposed. This finding is consistent with Schreuder and Baayen’s (1995) parallel dual-route activation model: transparent prefixed verbs are processed faster for both individual morphemes and whole word representations are activated unlike in opaque prefixed verbs, when only a whole word representation is activated as the prefixed root must be accessed to retrieve the representation.

To the best of our knowledge, our study represents the first attempt to address the differences in processing of a) derivation and inflection, b) opaque relative to transparent prefixed verb forms, and c) presumably inflective vs. derivational aspectual affixes in the highly inflectional Czech language.


LESS ISN’T ALWAYS MORE: THE NEGATIVE EFFECT OF PRIOR EXPOSURE IN NON-ADJACENT DEPENDENCY-LEARNING IN ADULTS
Ileana Grama, Annemarie Kerkhoff, Frank Wijnen (Utrecht University)
Contact: I.C.Grama@uu.nl

The existence of a critical period for language learning has been explained by the gradual development of working memory in infancy (the Less-is-More Hypothesis, Newport, 1988; Newport, 1990; Elman, 1993). Because infants start out with a limited working memory capacity they are forced to focus on the smaller units of language, becoming familiarized with them by the time they are able to process larger chunks of input and discover the relationships between these units. According to the Less-is-More Hypothesis, learning the rules of a language should be aided by the initial discovery of the component units. An important rule that learners of a language have to discover is the correspondence between non-adjacent grammatical morphemes in a sentence, marking for instance agreement or aspect (e.g. The boy’s holds the rabbit, The boy is holding the rabbit). It has been proposed that statistical learning enables learners to track co-occurrence statistics between dependent elements (Gómez, 2002; Gómez & Maye, 2005); the emergence of this non-adjacent dependency (NAD)-learning mechanism in infants appears to immediately precede sensitivity to morpho-syntactic dependencies (Santelmann & Jusczyk, 1998), and occurs well after sensitivity to independent grammatical morphemes (Shi & Lepage, 2008). In this study we investigate whether the acquisition of dependencies is aided by prior exposure to the dependent elements, in accordance with the Less-is-More Hypothesis.

Adult subjects listened to 15 minutes of a simple aXb ‘nonsense’ language, where a predicted b with 100% probability (familiarization); we tested learners’ sensitivity to the a_b dependencies through their grammaticality judgments of novel correct/incorrect aXb strings at test. We investigated the effect of prior exposure by introducing a one-trial pre-familiarization phase of varying lengths (1/2/8 minutes, Exp1-3), which consisted of aZ and Wb strings, such that the a/b tokens could be learned, but there was no indication of the a_b dependencies that would be heard in the subsequent familiarization phase.

Participants’ judgments of test items were less accurate the longer the pre-familiarization (Experiments 1-3), suggesting that, contrary to the Less-is-More Hypothesis, prior exposure had a detrimental effect on NAD-learning. Rendering the a/b tokens more acoustically salient (therefore easier to learn at pre-familiarization) also drastically reduced performance at test (Exp. 4). Lengthening the familiarization phase by 50% did not significantly improve learning (Exp. 5), suggesting that familiarity with a/b is catastrophic to learning a_b dependencies. Results are discussed in light of implications for the nature of NAD-learning and its role in acquisition: we suggest that dependency-learning must crucially operate before the dependent words are fully encoded, and that infant sensitivity to NADs should emerge at a very early stage in acquisition (Friederici et al., 2011).
Gelormini-Lezama & Almor (2011) showed that in Spanish, a null subject language, null pronouns are preferable to both overt pronouns and repeated names when referring to subject antecedents. Almor & al. (2014) showed similar preferences in Italian and even in Brazilian Portuguese, a language which is gradually losing the null subject. Here we ask whether in French, a non-null subject language, the same strong subject preference observed for null pronouns in null subject languages is observed for clitic pronouns (*il/elle*).

To answer this question, we used the paradigm of Gelormini-Lezama & Almor with French equivalents of their items. The French stressed form *lui* was chosen as the equivalent of the Spanish overt pronoun because both forms serve a contrastive/emphatic function in their respective language. The clitic pronoun *il* was chosen as the equivalent of the Spanish null pronoun because both forms are commonly used to refer to the most salient referent in the discourse. To test whether the French referential system parallels the Spanish system, we contrasted self-paced sentence-by-sentence reading times of French sentences with repeated name anaphors (2a), stressed pronouns (2b) and clitic pronouns (2c) referring to either subject (salient) or object (non-salient) proper name antecedents (1a). Analyses of second sentence mean reading times revealed a significant interaction between anaphor form and antecedent salience (F(1,41)=5.033, p<.01; F(2,70)=4.67, p<.05) such that stressed pronouns elicited longer reading time when referring to subject (2643 ms) than object (2395 ms) antecedents (F(1,41)=5.35, p<.05; F(2,70)=4.52, p<.04). Clitic pronouns showed the opposite pattern (1885 ms vs. 2096 ms) (F(1,41)=4.9, p<.05; F(2,70)=3.28, p=.078). No significant difference was found for repeated name (2059 ms vs. 1995 ms). Further comparisons revealed that, for subject antecedents, both repeated names and stressed pronouns elicited longer reading times than clitic pronouns (respectively, F(1,41)=4.52, p<.05; F(2,35)=3.08, p=.08 and F(1,41)=68.3, p<.0001; F(2,35)=66.83, p<.0001). Although French clitic pronouns show a subject preference similar to Spanish null pronouns, the preference for the subject antecedent was much stronger for Spanish zero pronouns (a 600 ms advantage in Gelormini-Lezama & Almor) than for French clitic pronouns (the 210 ms advantage we found here). This quantitative difference supports the idea that the preference for the syntactically salient antecedent is stronger in null subject than in non-null subject languages (e.g., Jegerski & al., 2011). Importantly, the object preferences for French stressed pronouns and Spanish overt pronouns were similar (248 ms vs. 107 ms, respectively). This suggests that while the Spanish overt pronoun and the French stressed form *lui* both serve a similar contrastive discourse function, the French clitic form is not as easy to process as a null pronoun for subject antecedents.

(1) **Sentence 1 - antecedent conditions:**
(a) Subject (salient) Jean a rencontré Maria. *John met with Maria.*
(b) Object (non-salient) Maria a rencontré Jean. *Maria met with John.*

(2) **Sentence 2 - anaphor form conditions:**
(a) Repeated Name Jean l’a trouvé triste. *John found her sad.*
(b) Stressed Pronoun Lui, il l’a trouvé triste. *HE found her sad.*
(c) Clitic Pronoun Il l’a trouvé triste. *He found her sad.*
In this study we investigate to what extent (i) definiteness (i.e., definite vs. indefinite noun phrases, NPs) and (ii) givenness (i.e., brand-new vs. linked) influence referential processing. We provide empirical evidence that ambiguous pronouns are more likely to be interpreted as referring back to a previously introduced referent, when this referent is realized as a definite noun phrase, regardless of its givenness.

Two main semantic approaches have been put forth to account for the distribution of definite and indefinite NPs. According to the first, a definite NP carries a familiarity presupposition (Kamp 1981, Heim 1982), whereas an indefinite NP indicates non-familiarity. According to the second, a definite NP presupposes its descriptive content and the uniqueness of its referent, while an indefinite NP only asserts the existence of its descriptive content and makes a non-uniqueness implicature (Russell 1905). Furthermore, definite and indefinite NPs introduce discourse referents that have different (cognitive) information statuses generally associated with different degrees of givenness: given, linked (also ‘bridged’ for definites), or brand-new (Prince, 1981; Strube & Hahn, 1999). Given NPs refer back to an explicitly introduced referent. Linked NPs require an inference and a certain level of relatedness between their anchor and the anaphor (Table 1 (2 a-b): a gym class context activates a script including the referents ‘instructor’ and ‘participant’). In contrast, brand-new NPs do not rely on an inference relation (Table 1 (2 c-d)).

Table 1: Sample experimental items (translated from German)

<table>
<thead>
<tr>
<th>(1) The gym class was crowded as usual.</th>
<th>(2) Michael observed ________</th>
<th>(3) When the door was closed, he put his towel on the mattress and started to stretch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) the committed instructor. [def+linked]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) a committed participant. [indef+linked]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) the American near the loudspeaker. [def+brand-new]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) an American near the loudspeaker. [indef+brand-new]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the semantic literature we will test two opposite predictions. According to the familiarity-based approach, we expect that both definiteness and givenness contribute to familiarity of the referent. Thus, [definite+linked] and [indefinite+linked] NPs should be more easily resolved towards a subsequent pronoun than [definite+brand-new] and [indefinite+brand-new]) NPs respectively. Alternatively, according to the uniqueness-based approach, we expect both types of definite and indefinite NPs to be resolved towards a subsequent pronoun, independently of givenness. Overall, we expect definite and indefinite NPs to contrast. In a visual world eye-tracking experiment participants (n=24) listened to short stories (Table 1, (1) – (3), n=48), while their eye fixations were monitored. All critical referents appeared in direct object (DO) position and were manipulated with respect to their morphological realization and information status, yielding four conditions (e.g. Table 1 (2 a-d)). Association strength to contexts was normed and was similar for referents in (2a) and (2b), ts < 0.3, ps > .7. The last sentence contained an ambiguous pronoun, he, which could refer back to the subject or the DO of the preceding sentence.

Results of separate linear mixed-effects models (0-1200ms; 0-400ms; 400-800ms; 800-1200ms) analysing proportions of looks to the DO over the subject revealed that, within 1200ms post pronoun onset, participants were more likely to interpret the pronoun as the preceding critical referent when that referent was definite rather than indefinite, t = 2.53, p = .011. Within the first 400ms, this main effect of Definiteness was accompanied by a marginal Definiteness X Givenness interaction, t = -1.93, p = .053, indicating that DO referents were initially more strongly fixated when they were definite and linked. We conclude that: (i) definiteness is a reliable predictor of how likely it is that a pronoun is interpreted as co-referring with a preceding NP and that (ii) givenness shows an early effect on pronoun resolution only for definite NPs. The results empirically sustain an analysis of definite NPs in terms of uniqueness rather than familiarity.
PRONOUN INTERPRETATION IS SENSITIVE TO VERB SEMANTICS, EXCEPT WHEN IT ISN’T: SELECTIVE VERB EFFECTS ON ITALIAN NULL AND OVERT PRONOUNS

Emily Fedele (University of Southern California) and Elsi Kaiser (University of Southern California)
emilyfedele@gmail.com

Prior work shows that anaphoric forms vary in their referential biases (e.g. Gundel et al. 1993). In Italian, according to Carminati’s (2002) syntactic approach, null pronouns refer to preverbal subjects, and overt pronouns refer to antecedents in other positions (e.g. objects). However, recent research suggests this may be an oversimplification (e.g. Serratrice 2007 on Italian), and a growing body of work suggests that verb semantics and coherence relations also play a key role (e.g. Kehler et al. 2008, Rohde 2008). Crucially, much of the work on semantic effects has focused on English pronouns, and it is not yet fully understood to what extent different pronoun forms (e.g. null/overt) in languages with richer anaphoric paradigms are influenced by semantic factors. We tested this for Italian.

We conducted two written sentence-continuation studies on Italian where we manipulated pronoun form (null/overt) and verb bias (implicit causality verbs with NP1/NP2 biases). In Exp1 (n=36), perché (‘because’) connected two clauses (ex.1). Exp2 (n=30) used the same stimuli, now separated into two sentences, with no connective (ex.2). (The auxiliary avere ‘have’ was used to show the presence of the null pronoun.) Continuations were analyzed to see if they are guided by verb semantics (NP1 verb=>sub, NP2 verb=>obj) or by grammatical biases of pronominal forms (null=>sub, overt=>obj), or both. (We also had a no-prompt condition where people could freely produce different forms; those results for null and overt are largely in line with what we find in the prompt conditions.)

(1a) Lo studente ha deluso_{NP1} / criticato_{NP2} lo chef perché lui ha…
   The student has disappointed_{NP1} / criticized_{NP2} the chef because he has… [overt]
(1b) The student has disappointed_{NP1} / criticized_{NP2} the chef because has… [null]

(2a) The student has disappointed_{NP1} / criticized_{NP2} the chef. He has … [null]
(2b) The student has disappointed_{NP1} / criticized_{NP2} the chef. Has… [null]

RESULTS EXP1 (‘because’): As expected, nulls show a stronger subject preference than overt (p’s<.05). Crucially, we also find a verb effect for both forms: NP1 verbs elicited mostly subject continuations with both null and overt pronouns (null: 85% sub, 11% obj; overt: 67% sub, 20% obj; p’s<.05); NP2 verbs elicited mostly object continuations with both (null: 34% sub, 60% obj; overt: 23% sub, 66% obj; p’s<.05). Although the verb effect is stronger with overt than with nulls (formXverb interaction), it is significant for both forms. Thus, contrary to the common view that nulls refer to subjects and overt to non-subjects, we find once verb semantics are taken into account, nulls can refer to objects and overt to subjects. The semantic effects are stronger with overt verbs, especially subject continuations (NP1: 90%; NP2: 71%). However, overt pronouns flip from object to subject based on verb bias, like Exp1: NP1 verbs elicited more subject continuations (63%), NP2 verbs more object continuations (63%; p<.05). There is a significant effect of ‘experiment’ (Exp1 vs2) for nulls but not overt verbs: The verb effect persists for overt but weakens for nulls in Exp2.

Our comparison of intra- vs. inter-sentential configurations shows that prior characterizations of null vs. overt are insufficient, and neither a purely form-based nor a purely verb-based approach is enough. We argue for a theory that can capture effects of verb semantics and discourse structure for different types of referring expressions.
NP STATUS IN THE ESTABLISHMENT OF FOCUS AND PROCESSING OF ANADEIXIS

Derya Cokal (University of South Carolina), Patrick Sturt (University of Edinburgh) & Fernanda Ferreira (University of South Carolina)
cokal@mailbox.sc.edu

Different referential expressions can be sensitive to different factors (6, 5). It, this and that have been argued to signal different procedural instructions as markers of the saliency level of their referents in the addressee’s memory (2, 3, 7). It serves to maintain attention, whereas this/that direct readers’ attention from the focused entity to a less focused entity (3, 5, 7, 8). The role of noun phrase (NP) status in the establishment of focus of pronominal it/this/that has received little attention. Our predictions were: (a) references with it to the focused NP would lead to shorter fixations than with this; (b) references to the less focused NP with this/that would lead to shorter fixations than with it.

We performed two eye-tracking reading experiments. In Experiment 1 (40 items, 40 Native English speakers), we manipulated the referents of it, this and that (e.g. the room or a jug) by mismatching/matching referential expressions with the features of NPs in subject/object positions (e.g. a window or handle). In 1a, it, this and that referred to the room, in the subject position, and in 1b, they referred to the jug, in the object position (see below). Experiment 1: The room was small and had a large jug in the centre.

1a) It/this/that had a large window and looked stylish.
1b) It/this/that had a large handle and looked stylish.

- In 1a, the fixations in the critical regions (i.e. window/handle) of second and total pass reading times were longer when this referred to the NP in the subject position than when it referred to the subject NP (ps=.046). This preferred the object NP references.
- In 1a and 1b, there were no object and subject preferences for it.
- In 1a and 1b, fixation times for that did not differ between subject and object reference, but were longer in both cases than for this and it.

In Experiment 2 (40 items, 40 Native English speakers), it/this/that were given in the object position to explore whether their presentation in different structural conditions would affect their NP preferences. Here, the verb put created a focus and made the first NP (e.g. wine glass) salient, whereas the preposition made the second NP (e.g. the bottle) less salient (see below). Experiment 2: Joseph put the wine glass next to the bottle. Before washing up, he grasped it/this/that 2a) by its stem and put it on the sideboard.

2b) by its cork and put it on the sideboard.

Again, we manipulated their referents by mismatching/matching referential expressions (e.g. wine glass or bottle) with features of the first/second NPs (e.g. stem/cork).

- In 2a, references to the focused NP (e.g. wine glass) with it led to shorter fixations than references with this in the critical regions (i.e. stem/cork) of second pass reading times (ps=.025).
- In 2b, references to the less focused NP (e.g. the bottle) with this led to shorter fixations than references with it in the critical regions of second pass reading times (ps=.049).
- In 2a and 2b, fixation times for that did not differ between subject and object reference, but were longer than this and it.

Our results show that this and it were sensitive to different NP statuses in search of referent processing and establishment of focus. That does not share the same sensitivity to the NP status. This is contrary to the assumptions on the role of NP status in the establishment of the focus of that in 3, 4, and 7. In line with 6 and 1, we argue that a saliency-based account to establish focus is possible for this and it, but is not generalizable to that.

Understanding how people comprehend non-literal language has received much attention in the psycholinguistic literature [1,2] with non-literal (i.e., indirect) replies often being used as a face saving tool [2]. Although there is evidence suggesting that indirect meaning may be processed differently from literal meaning [3], the temporal aspects of indirect meaning comprehension remain uncertain. Using an Eyelink 1000, we monitored the eye movements of 36 participants reading 48 experimental vignettes (see example below). We manipulated context to be either positive, negative, or neutral. The target region (underlined) stayed constant across the three versions of any one item. For example:

Ken and Maxine are university students and friends. Last Friday, Maxine went on a date (and it went well/terribly/…). Ken asked “How did the date go?” She replied “Apparently I am not his type.” Maxine had been single for nearly a year.

The indirect meaning interpretation of the target region is only incongruent in the positive condition. What is critical is when in the eye-movement record we find evidence that readers are sensitive to this (in)congruency. If the literal and non-literal meanings are constructed rapidly in parallel, we would expect to see disruption to the eye-movement record arising at an early point (i.e., on early measures of processing associated with the target sentence) when this sentence appears in the positive context (i.e., where a negative meaning is anomalous given prior context). However, if the literal meaning is constructed first and then discarded before the non-literal interpretation is constructed, then we would expect to find delayed effects related to this anomaly.

We found clear disruption in the eye-movement record on reading of the final sentence (and none on the target region itself). This was most clearly seen as an increase in regression path reading times and as an increase in regressions out of the final sentence in the positive condition, relative to the negative condition. These findings provide strong evidence that processing disruption arising from the identification of a relevance violation and processing of negative indirect meaning occurs only after a delay. This is compatible with the view that the literal interpretation was computed before the non-literal interpretation. This stands in contrast to other research on non-literal language processing which has shown evidence of more rapid comprehension of non-literal meaning. We propose that the additional face-saving function of the non-conventional indirect replies we examined takes up processing resources resulting in delayed comprehension of the indirect meaning itself.

References
Predicting what will be mentioned in utterances can be an easy task as long as the preceding context is restrictive (Altmann & Kamide, 1999). Some linguistic processes seem automatic and unintentional; for example, people show biased looks to objects that are related to pre-activated words even though they are implausible in the context (Rommers, Meyer, Praamstra, & Huettig, 2013). To investigate making prediction is also such an automatic consequence of comprehension processes, we manipulated language proficiency (native or L2 speaker of English) and available resources (prediction with or without an additional memory task) in a series of Visual World eye-tracking experiments.

In Experiment 1, 24 L1 and 24 L2 speakers of English listened to sentences such as “The woman will fold/ find the scarf.” while viewing pictures containing a target object (scarf) and distractor objects (violin, high-heels, and piano). The target object was either predictable or unpredictable depending on the semantic restriction of the preceding verb in the spoken sentence. In Experiment 2, another set of 24 L1 and 24 L2 participants were shown 5 words before viewing each picture, and recalled them after listening to the sentences. This memory task demanded participants to use cognitive resources, and if predictive processing is automatic, it would remain unaffected by the memory load.

When there was no memory task (Experiment 1), both L1 and L2 speakers were similarly able to predict target objects, and directed their eyes towards the objects about 1000ms before the acoustic onset of the target words. In contrast, when there was a memory load (Experiment 2), it was about 300ms before the target word onset that L1 speakers started to show bias towards predictable objects, and such bias occurred only about 150ms after the target onset in L2 speakers. Comparing the Experiment 2 to the Experiment 1, the prediction effect was delayed in L1 speakers, and was eliminated in L2 speakers. There was no difference in the word-recall accuracy between the L1 and L2 speakers (73% vs. 68%).

The weaker prediction under the working memory load (i.e., when less resources were available) suggests that cognitive resources are required for prediction or for processing associated with prediction (e.g., indexing the objects). Furthermore, as online comprehension of L2 is generally more demanding in terms of cognitive load than L1 (Ardila, 2003), the greater interference effect of the memory load on L2 speakers suggests that the resources used for prediction also underlies general comprehension abilities.

References

Mechanism under Auditory Perceptual Simulation: Speech Speed or Attitude Biases?  
Peiyun Zhou (University of Illinois at Urbana-Champaign & Beckman Institute) & Kiel Christianson (University of Illinois at Urbana-Champaign & Beckman Institute)  
pzhou5@illinois.edu

**Introduction:** Embodied cognition theory proposes that semantic meaning is derived from the activation of perceptual features associated with physical experiences and sensations (e.g. Barsalou, 1999; Zwaan, 2008). Two eye-tracking studies demonstrated that readers' eye movements can be affected when they mentally simulate characteristics of voices or speakers depicted in texts (Stites, Luke, & Christianson, 2013; Yao & Scheepers, 2013). This phenomenon has been referred to as Auditory Perceptual Simulation (APS). A recent eye-tracking study (Zhou & Christianson, submitted) found that APS of native and non-native speech impacts readers' sentence processing: readers' reading speeds were modulated by the speed of the native (faster) or non-native speaker (slower) that they were perceptually simulating. Moreover, APS of either native or non-native speech was observed to both increase reading speed and improve comprehension during silent reading compared to normal silent reading without APS. College students showed 20% improvement in accuracy in reading morphsyntactically complex and semantically implausible sentences (1d). However, it might be argued that the slower reading speed when activating APS of non-native speech in Zhou and Christianson's study was triggered by readers' biases towards non-native speech, since social attractiveness studies have shown that listeners have more negative attitudes towards non-native than native speech (Edward, 1977; Lev-Ari & Keysar, 2010). To investigate whether the APS effects observed by Zhou & Christianson were triggered by the native and non-native “speakers’” speech rates or the readers’ biases towards native and non-native speech, an eye-tracking experiment was conducted with manipulation of a faster Hindi-accent speech and a slower native American English speech.

**Method:** To induce participants (N=80) to perceptually simulate the speech of a native and non-native English speaker, examples of each speaker's voice were provided. One American English and one Indian non-native English speaker were recorded reading four passages (500 words each). These recordings, which differed only in speaker accent and rate, were played for participants as they were shown a photograph of each speaker. Two speakers' photos and name recordings were used (counterbalanced) prior to each sentence in order to cue the "speaker's" voice during sentence processing. 48 target sentences and 96 fillers were read silently by participants while their eyes were tracked. The target sentences manipulated structural and semantic characteristics that have been shown to result in variable processing and comprehension difficulty (1). This resulted in a 2(structure) x 2(plausibility) x 2 ("speaker") fully factorial design. After each sentence, participants answered a paraphrase verification question (1e). Half-way through the experiment, two more recordings of the native and non-native speakers' voices were played to remind participants what they sounded like. Eye movements were monitored using an EyeLink 1000 desktop eye tracker. In addition, a social attractiveness scale questionnaire was administered to measure participants' attitudes toward the "speakers"' speech.

1. a. “The bird that ate the worm was small.” (SRC-plausible)  
   b. “The worm that ate the bird was small.” (SRC-implausible)  
   c. “The worm that the bird ate was small.” (ORC-plausible)  
   d. “The bird that the worm ate was small.” (ORC-implausible)  
   e. The worm/bird ate the bird/worm. The bird/worm was small. (True/ False)

**Results:** Online data demonstrated that APS of the faster non-native English speech led to faster reading speed than APS of the slower native speech, while survey results showed that participants still had more negative attitudes towards the non-native speech, consistent with Zhou and Christianson’s study. The results suggested that APS effects were triggered by the speech rate of the native and non-native speakers instead of sociolinguistic biases towards the native and non-native speech. Moreover, plausibility and syntactic structure significantly affected the sentence reading time (implausible> plausible; ORC>SRC) and response accuracy (plausible> implausible; SRC>ORC).
RESUMPTIVE PRONOUNS AMELIORATE ISLAND VIOLATIONS IN FORCED-CHOICE TASKS
Lauren Ackerman, Michael Frazier & Masaya Yoshida (Northwestern U)
Imackerman@u.northwestern.edu

Introduction: The syntax literature reports that resumptive pronouns (RPs) can ameliorate island violations [1, 2]. However, much psycholinguistic literature finds RPs no more acceptable than equivalent gaps, [3, 4, 5]; Cf. [6], possibly due to methodological differences. In psycholinguistics studies, RP and illicit gap conditions are not typically shown together, so participants do not directly compare the two conditions. The contrasts reported in the syntax literature come from direct comparison of minimal-pair sentences. Moreover, production tasks consistently elicit RPs in islands [7, 8]. Explanations for this discrepancy include the following. Formal acceptability tasks may not detect amelioration effects from RPs because RPs do not ameliorate island violations, Cf. [9]. Thus their presence in speech may be an artifact of production. Alternatively, amelioration effects may be too subtle for detection without direct comparison, thus the production-comprehension split may be due to task type.

Experiments: Six forced-choice tasks on Amazon Mechanical Turk crossed three island types (1: relative clause island, 2: adjunct island, 3: wh island) with two task types (forced-choice sentence completion, SC; full-sentence forced-choice, FS). In SC tasks, sentences appeared with a blank in the position of the trace and preceding verb (underlined below). In the FS tasks, the same stimuli were presented in Gap-RP pairs, without blanks or underlines.

(1) a. Which man did Jane say that the parent scolded (him) forgave the girl’s mistake?
   b. Which man did Jane say that the parent who scolded the caregiver forgave (him)?

(2) a. Which man did Jane say that, after the parent scolded (him), the caregiver forgave the
girl’s mistake?
   b. Which man did Jane say that, after the parent scolded (him), the caregiver forgave
   (him)?

(3) a. Which man did Jane ask when the parent scolded (him)?
   b. Which man did Jane say that the parent scolded (him)?

In each experiment, 40 participants indicated the more natural option. The design was 1x2: Location (Island vs Nonisland). If RPs can ameliorate island violations, RPs should be selected more often than Gaps in the Island conditions (a). If RPs cannot ameliorate islands, both options should be chosen at chance in islands, since both are equally unacceptable. In either scenario, the licit gap option should be preferred in nonisland conditions (b). If the discrepancy comes from a difference in production and comprehension mechanisms, we may expect the above predictions to obtain only in the SC task. Alternatively, if previous production/comprehension studies probe different mechanisms, similar results should obtain throughout. Using binomial generalized linear mixed models, all experiments show a main effect of Location, with an RP preference in island conditions (is), and a Gap preference in nonisland conditions (ni) (all cells p<0.0001):

<table>
<thead>
<tr>
<th></th>
<th>SC</th>
<th>Adjunct islands</th>
<th>Wh islands</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC islands</td>
<td>RP_{is}: 82%, RP_{ni}: 38%</td>
<td>RP_{is}: 80%, RP_{ni}: 54%</td>
<td>RP_{is}: 60%, RP_{ni}: 11%</td>
</tr>
<tr>
<td>FS</td>
<td>RP_{is}: 80%, RP_{ni}: 25%</td>
<td>RP_{is}: 79%, RP_{ni}: 57%</td>
<td>RP_{is}: 63%, RP_{ni}: 14%</td>
</tr>
</tbody>
</table>

Discussion: First, a gap preference in nonisland conditions suggests that RPs are sensitive to grammatical structure: RPs are dispreferred when a gap is licit. Second, a strong RP preference in islands suggests RPs do ameliorate, if not rescue, island violations. Third, that this effect obtains in SC and FS tasks suggests that production/comprehension mechanisms are not the source of the discrepancy. Unexpectedly, there is a substantial number of RP completions in the nonisland conditions. This is may be due to the complexity of the sentences, which could be a source of increased acceptability for RPs [10].

Conclusion: Our study indicates that in a direct choice between an RP and an illicit gap, amelioration effects obtain. Furthermore, the production-comprehension split may be due to task type, rather than a difference in mechanisms. If so, the methods by which acceptability is measured may need to be revisited for similar, subtle effects.

JUDGING PARADIGMATIC RELATIONS: A COLLECTION OF RATINGS FOR ENGLISH
Gabriella Lapesa (University of Osnabrück, University of Stuttgart), Sabine Schulte im Walde (University of Stuttgart), Stefan Evert (FAU Erlangen-Nürnberg)
glapesa@uos.de

1. Goal and Motivation. We introduce a collection of semantically related English word pairs, rated for the strength of the semantic relation holding between them. Our collection presents several elements of novelty with respect to comparable datasets such as WS353 (Finkelstein et al. 2002) and RG65 (Rubenstein and Goodenough, 1965).

- While other datasets focus on similarity (RG65) or similarity/relatedness (WS353), we tackle three different paradigmatic relations, namely synonymy, antonymy, hypernymy.
- While other datasets (WS353 and RG65) only contain ratings for nouns, we collected ratings for three different parts of speech: nouns, verbs, adjectives. The comparison across parts of speech is expected to highlight conceptual differences of the three relations across word classes. For example, the concept of hypernymy has been widely investigated with respect to nouns, but little attention has been devoted to its application to verbs.
- While other datasets (WS353 and RG65) do not consider gradedness in similarity, we collected ratings at three different degrees of relatedness: strongly, weakly and negatively related. Examples of related words for the target artist in the synonymy relation, are: painter (strongly related), creator (weakly related), scientist (negatively related). As negatively related, we chose antonyms for synonyms and hypernyms, synonyms for antonyms.
- While other datasets (WS353 and RG65) collect ratings uni-directionally, we introduce directionality as a parameter in the collection of the ratings. For every ⟨target,relation,relatum⟩ triple, we collected forward and backward ratings (e.g., artist-synonym-painter vs. painter-synonym-artist). The question at issue is to which extent asymmetry affects relations (e.g., synonyms vs. antonyms) with regard to parts of speech (verbs/adjectives vs. nouns).

2. Target selection & Collection features. The target selection was conducted in a two-step process. First, a generation experiment asked native speakers to generate related words (synonyms, antonyms and hypernyms) for 99 English targets per part-of-speech. The targets for this generation experiment were chosen using the stratified sampling technique by Scheible and Schulte im Walde (2014), that relies on a random selection from WordNet balanced for a) target frequency classes, b) polysemy classes, and c) the WordNet semantic class. The generation experiment was carried out by Giulia Benotto and Alessandro Lenci at the Computational Linguistics Lab, University of Pisa. Second, from the generated word pairs, we selected the ⟨target,relation,relatum⟩ triples to be rated. We identified ⟨target,relation⟩ pairs (e.g., ⟨artist, synonym⟩) such that a) at least 2 different relata had been produced in the generation experiment; b) the strongly related word (e.g., painter) had been produced at least 4 times; c) the weakly related word (e.g., creator) had been produced twice (preferred) or once; d) the negatively related word had been produced at least twice for the opposing relation (e.g., ⟨painter, antonym, scientist⟩). In total, we selected 284 targets and collected ratings for 1,704 target / relation / related word / direction combinations, on a scale from 0 (not related) to 5 (fully related). Ratings were collected using Amazon Mechanical Turk. Every ⟨target,relation,relatum⟩ triple was rated by 10 subjects for each direction.

3. What we will present. In the presentation, we will a) provide more details of the selection procedure; b) report on quantitative and qualitative analyses of the collection; c) describe corpus-based modeling of the ratings.

As comprehenders hear speech or read text, they anticipate or predict upcoming words (e.g., Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2003, 2004). How broad are these predictions, and how many predictions do comprehenders make for an upcoming word? If a listener who hears the sentence fragment “After doing his laundry, Mark always seemed to be missing one…” strongly expects to hear the word “sock” next, is the word “shirt” partially expected as well, is it actively inhibited, or is it ignored?

The present research assessed the scope of prediction by measuring its effects on the processing of subsequently presented stimuli using the cumulative semantic interference paradigm. Across three experiments, 220 subjects each named 94 pictures (60 critical). The critical pictures constituted 12 semantic categories of five pictures each (articles of clothing, animals, etc.). Half of the pictures were presented in isolation; the other half were preceded by high-cloze sentences presented via RSVP with the last word omitted (see above for an example).

Despite differences across experiments in the distribution of conditions within each category and the probability that the sentence fragment matched the subsequent picture, every experiment showed the same pattern of results. Naming sock in isolation slowed the subsequent naming of the picture shirt in isolation – the standard cumulative semantic interference effect (represented by the positive slope of the black line in the Figure). Critically, although picture naming was nearly 200 ms faster after sentences, the interference effect was not modulated by the context (bare vs. sentence) in which either picture was presented (represented by equivalent slopes for the two lines). In other words, producing “sock” slowed the subsequent production of “shirt” by the same amount regardless of whether zero, one, or both pictures were presented after highly constraining sentences.

According to the only model of cumulative semantic interference that can account for such a pattern of data (Oppenheim, Dell, & Schwartz, 2010), this indicates that a sentence pre-activates its best completion (sock) but does not affect the activation of less likely completions (shirt), as any modulation of non-target activation would have affected the amount of interference for picture naming after sentences. Thus, the present research suggests that comprehenders pre-activate only the most probable completion for each sentence.
LEXICAL REPRESENTATIONS OF NOUNS IN GERMAN RELY ON UNDERSPECIFIED GENDER FEATURES
Andreas Opitz (Leipzig University) & Thomas Pechmann (Leipzig University)
andreas.opitz@uni-leipzig.de

Traditionally, instances of a certain grammatical category (e.g., gender or case) are categorically labeled to differentiate between distinct classes (e.g., masculine, feminine, neuter referring to gender in German). Current morphological theories, however, propose more fine-grained analyses of these categories. Inflectional markers that represent (or realize) one of those categories may either clearly correspond to one particular syntactic context, or may be ambiguous by referring to more than one syntactic context, i.e., an inflectional paradigm may exhibit instances of syncretism. This phenomenon is accounted for by abstract feature decomposition and the concept of underspecification in almost all current morphological frameworks. The over-all idea behind these two concepts is a decomposition of 'traditional' labels of morphosyntactic categories into more abstract, binary features, thereby yielding the possibility to refer to natural classes of such categories. Thus, the three instances of grammatical gender in German could be described by two abstract binary features [±f] and [±m]: ‘feminine’ [+f, -m], ‘masculine’ [-f, +m], ‘neuter’ [-f, -m]. Interestingly, in all recent accounts of German inflection (Blevins, 1995, see also Opitz et al., 2013) that make use of feature decomposition and underspecification no reference to any [-m] feature value is necessary to derive the system of inflectional markings. As a consequence one may ask if such a feature value [-m] that is never used for morphosyntactic operations is represented in the system at all. Thus, the basic question of our study was whether lexical representations of inherent features, i.e. gender information of nouns, are fully specified or rather underspecified.

The question whether morphological processing relies on more traditional or more fine-grained, abstract categories is barely addressed in psycholinguistic research yet. Although there is first evidence that underspecification and more abstract feature-decomposition are involved in human language processing (see Clahsen et al., 2001, for behavioral studies and Opitz et al., 2013, for ERP evidence) all of the studies so far addressed the processing of inflectional morphology (affixation and stem alternations). In contrast, the lexical representation of uninflected stems is regarded as fully specified with respect to inherent features. Our findings from a series of experiments, however, challenge this view. We obtained results indicating that the processing of nouns in German, although matched for length, frequency and morphological complexity, differs systematically between grammatical genders. Using different behavioral tasks (lexical decision, gender decision, and gender verification) we found that masculine nouns yielded longer latencies than feminine nouns which are approximately of equal type frequency. These findings are interpreted as reflecting differently specific lexical representations of gender features for German nouns, i.e. feature sets that differ in their amount of informational content. Thus, we propose specifications of gender features in the lexical representation of nouns in German that reflect the observed processing differences and are also compatible with morphological analyses that go without any feature [-m]: two features [+m, -f] for masculine nouns, as opposed to one feature [+f] for feminine nouns, and [-f] for neuter nouns.

References
SUBLIMINAL MORPHOLOGICAL SPEECH PRIMING

Roberto Bottini (University of Milan-Bicocca), Davide Crepaldi (University of Milan-Bicocca) roberto.bottini@unimib.it

To what extent language can be processed unconsciously? The processing architecture of unconscious word recognition is a prolific field in psycholinguistics. Nevertheless, whereas hundreds of studies have been devoted to unconscious word processing in the visual modality, similar studies in the auditory modality can be counted on the fingers of one hand. One reason for this discrepancy is methodological. The classic paradigm used to investigate the early stages of word recognition is masked priming, and it was developed for the visual modality. In a standard masked priming experiment, a lowercase prime is presented briefly (e.g. 40ms) sandwiched between a forward mask (e.g. XXXX) and an uppercase target. Although participants are usually not aware of their presence, unconscious primes facilitate the processing of the target words at the lexical, morphological and semantic level.

Due to the characteristics of spoken words, developing a masked priming paradigm in the auditory domain has constituted for years a methodological challenge. Recently, Kouider and Dupoux (Kouider, S. & Dupoux, E., 2006, Subliminal Speech Priming, Psych.Sci., 16, 617-25) suggested a way to exit the impasse. In their experiment a spoken word was time-compressed and hidden within a stream of speech-like masks with similar spectral characteristics. The target was played right after the compressed prime at a normal speech rate and it was superimposed over the backward masks. The authors found that lexical decision on the target was facilitated when prime and target were the same word (repetition priming) even though participants were not aware of the presence of the prime (as indicated by low d-prime values in a subsequent prime recognition test). This result suggested that spoken language can be processed unconsciously up to word form at the least.

In this study we tested whether, as in the visual domain, the unconscious processing of spoken words extends to their morphological structure. The same stimuli were used both in a visual version and an auditory version of a masked prime paradigm, allowing for the first time a direct comparison between the unconscious processing of written and spoken words. In the experimental trials primes were related to the target by identity (deal - DEAL), or morphological derivation (dealer - DEAL). The priming effect was measured against matched unrelated pairs (e.g., string - DEAL; gardener – DEAL). We used Kouider & Dupoux methodology in the auditory version of the experiment, with prime compressed down to 50%, 40% and 35% of their original duration. Participants were randomly assigned to one of the three compression rates (CR). In the visual version primes were displayed for 35ms. In both versions the task was a lexical decision on the target words. Moreover, participants performed a prime recognition test after the experiment in which they were asked to perform a lexical decision on the primes.

Results of the prime recognition test revealed that primes were supraliminal at 50% CR, but were effectively masked at 35% and 40% CR. Not surprisingly, RT analysis showed both repetition priming and morphological priming at 50% CR, with supraliminal primes. Crucially, the same result was observed at 40% CR with subliminal primes. At 35% CR neither repetition nor morphological priming was observed, probably due to excessive degradation of the prime. Additional analysis were preformed for the 40% CR in order to confirm the genuinely subliminal nature of the priming effect: After excluding participants that were better than chance in the prime detection task, or trials in which the prime was correctly recognized by each subjects, results did not change.

The outcome of the visual version of the experiment confirmed previous findings showing both repetition priming and morphological priming effects. These results suggest that unconsiously perceived spoken words are processed up to their morphological structure mirroring the level of processing found for written words. This opens interesting possibilities for the exploration of the similarity between visual and auditory word identification.
EFFECTS OF TALKER AND NOISE VARIATION ON RECOGNITION MEMORY FOR
SPOKEN WORDS
Angela Cooper (Northwestern University) & Ann R. Bradlow (Northwestern University)
akcooper@u.northwestern.edu

Prior research has demonstrated that listeners are sensitive to changes in the
indexical characteristics of speech input, suggesting that these features are integrally
encoded in memory for spoken words (e.g. [1]). The integrality of linguistic and indexical
information may be at least partially accounted for by the fact that the acoustic correlates of
these features are inherent to the speech signal itself, stemming from the same sound
source. However, given that listeners frequently must contend with concurrent environmental
noise, the question remains as to the extent to which listeners encode contextual details
external to the speech sound source. While recent work has provided initial evidence that we
store integrated representations of simultaneously heard linguistic and non-linguistic auditory
signals [2], we provide a stronger test of this hypothesis by utilizing spectrally-separated
speech and noise signals, such that changes in a simultaneously presented non-speech
signal (background noise) from exposure to test would not be accompanied by concomitant
changes in the target speech signal.

For the current study, we assessed listeners’ explicit memory for spoken English
words as a function of consistency versus variation in the talker’s voice (talker condition) and
the background noise (noise condition) using a recognition memory paradigm. English
listeners first completed a word identification task, where they were exposed to a set of
items that were either produced by two female talkers (talker group) or combined with two
types of noise (noise group). This was followed by a recognition memory task, where half of
the items were the same exemplars provided in the word identification task and the other
half were different exemplars, with either a change in talker or noise. For each item,
participants were asked to recall whether or not they had heard it in the first task. Higher
accuracy rates for same-exemplar repetitions, where the talker or noise was consistent from
exposure to test, relative to different-exemplar repetitions would suggest that listeners
encoded context-specific information alongside linguistic information in memory. We also
manipulated the lexical frequency of the items to test the prediction that specificity effects
are stronger in low relative to high frequency words [3]. For the noise condition, to generate
spectrally-separated items, two sets of stimuli were created by combining speech stimuli that
had been low-pass filtered at 5 kHz with narrow band-pass filtered white noise from 7-10
kHz for one set and with a 6 kHz pure tone for the other set.

Consistent with previous work (e.g. [1]), hit rates revealed that listeners were
significantly more accurate at recalling same-talker repetitions relative to different-talker
repetitions. Furthermore, extending this prior work, hit rates for repetitions with the same
background noise were significantly higher than for repetitions with different noise. Finally,
lexical frequency was not found to significantly mediate the strength of the exemplar
specificity effect in either condition. These findings indicate that listeners encode both within-
signal speech features (e.g. talker identity) as well as speech-extrinsic information (e.g.
background noise) into integrated cognitive representations, even when the two auditory
streams are spectrally non-overlapping.

recognition memory. *Journal of Experimental Psychology: Learning, Memory and
Cognition*, 22, 1166-83.


*Psychological Review*, 105, 251-279.
Previous research has shown that understanding action-related (Meteyard, et al., 2007) and spatial (Richardson, et al., 2003) words involves simulating associated sensorimotor experiences. The current study examined the ability of verbs’ implicit spatial biases to displace visual attention in vertical and horizontal spaces. Participants read individually presented verbs derived from Meteyard & Vigliocco (2009) and completed a version of Posner’s (1980) visual probe detection task. The spatial congruency between the verb’s implicit spatial bias and the visual probe location was consistently manipulated. In Study 1, verbs had a vertical bias (upward or downward) and the visual probes were presented above or below the central fixation. In Study 2, the verbs had a horizontal bias (rightward or leftward) and the probes were presented left or right of the central fixation. Analysis of the reaction times indicated a reliable visual target effect in Study 1, with faster reaction times for upward targets (cf. Goldring & Fischer, 1997). More importantly, there was a reliable interaction between verb bias and probe location in Study 2, with faster responses for visual probes congruent with the verb’s bias. Our findings suggest that verbs’ implicit spatial biases are simulated during language comprehension and that this simulation leads to the displacement of visual attention, at least across the horizontal space. Our findings provide further support to the grounded theories of cognition.
The syntactic mismatch between antecedent and ellipsis in a verb phrase ellipsis (VPE) sentence like *The problem was looked into by Kim just like Lee did* makes it unacceptable. Some analyses claim that such sentences are ungrammatical because syntactic reconstruction is required for an elided VP. However, there are examples of relatively acceptable verb phrase ellipses (VPE) in which the antecedent syntactically mismatches the ellipsis, e.g., *The problem was to have been looked into, but obviously nobody did*. Kehler (2002) accounts for the differences in acceptability in terms of how different discourse relations (e.g., causality, resemblance) impact the processing of various linguistic constructions. Kehler proposes that mismatching instances of VPE are not ungrammatical, and are acceptable when the relation between the clauses is *causality*, as in the ‘looked into’ sentence. However, when the relation is *resemblance*, syntactic similarity is required to easily form a semantic interpretation, resulting in unacceptability.

Frazier and Clifton (2006) tested Kehler’s prediction that mismatching VPEs will be relatively acceptable when they express a causal relation between the clauses, and found that a mismatch impaired causal and resemblance VPEs equally. However, their manipulation of type of relation involved only the connective between clauses: *because* or *even though* for a causal relation, *just like* for resemblance. As Kehler (2002) makes clear, the form of the connective does not guarantee the intended interpretation.

To correct this shortcoming, two experiments were conducted in which causal and resemblance sentences were presented as replies to a question about the reason for an event or a question about who was similarly involved in an event. The experiments used the same 16 items, except that Experiment 2 added a "too" to the end of the causal sentences. An example of one item appears in (1):

1. **CAUSAL/Mismatch**: Why did Lee get involved in checking out the problem, do you think? The problem was looked into by Kim so Lee did (too).
   
2. **CAUSAL/Match**: Why did Lee get involved in checking out the problem, do you think? Kim looked into the problem so Lee did (too).
   
3. **RESEMBLANCE/Mismatch**: Who else besides Lee got involved in checking out the problem, do you think? The problem was looked into by Kim just like Lee did.
   
4. **RESEMBLANCE/Match**: Who else besides Lee got involved in checking out the problem, do you think? Kim looked into the problem just like Lee did.

In Experiment 1, 64 undergraduates rated 16 examples of these replies on a 7-point scale from "unnatural" to "natural", in a fully counterbalanced design. In Experiment 2, 32 Mechanical Turk "Turkers" did the same thing. (Different filler items were used, changing the overall ratings.) The results, shown in Table 1, clearly support Frazier and Clifton's (2006) conclusions. In each experiment, mismatching VPEs were rated significantly lower than matching examples (t > 4.0) and the interaction of discourse relation and match was nonsignificant (t < 1.2). Mismatched VPEs are unacceptable, regardless of the discourse relation they express.

<table>
<thead>
<tr>
<th>Expt</th>
<th>CAUSAL/MM</th>
<th>CAUSAL/MATCH</th>
<th>RESEMB/MM</th>
<th>RESEMB/MATCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.16</td>
<td>2.60</td>
<td>2.45</td>
<td>3.10</td>
</tr>
<tr>
<td>2</td>
<td>3.75</td>
<td>4.77</td>
<td>3.93</td>
<td>5.14</td>
</tr>
</tbody>
</table>

This conclusion, to be sure, leaves open the question of why some mismatched VPE sentences are relatively acceptable. One possibility was suggested by Grant et al. (2012), who showed the presence of a modal phrase expressing the possible but uncertain actuality of the antecedent event reduced the unacceptability of mismatching VPEs. This uncertainty (which was characteristic of nearly all of Kehler’s, 2002, examples of causal ellipses) may have led readers to expect a second clause addressing it, facilitating the comprehension of the ellipsis and resulting in relative acceptability of an ungrammatical sentence.
DIVERGING ONLINE STORAGE COSTS FOR WH-ADJUNCTS ARE PREDICTED BY
THE GRAMMATICAL THEORY
Arthur Stepanov (University of Nova Gorica) & Calum Riach (University of Nova Gorica)
arthur.stepanov@ung.si

The problem: It has been argued that there is an online storage cost associated with an incomplete wh-filler-gap dependency reflected in increased response times over an extended region between the filler and the gap in self-paced reading (Chen, Gibson & Wolf 2004). The relevant studies traditionally concentrated on processing wh-subjects and objects and rarely mention wh-adjuncts (e.g. when, why). Recent advances in syntactic theory indicate that wh-adjuncts manifest a rather different syntactic behavior than wh-arguments, so it is reasonable to ask whether this difference is also reflected in online behavioral measures. In particular, Rizzi (1999), Stepanov & Tsai (2008), a.o., have argued that why does not trigger a wh-dependency in the same way other wh-items do. In other words, why is base-generated in its surface position in the left periphery of the clause (=Comp), unlike other wh-adjuncts which originate within a predicate phrase they modify and then are displaced to Comp in overt wh-movement languages. If wh-dependencies are associated with online storage costs, we may expect why not to trigger the same kind of costs as other wh-adjuncts. In the present study, we asked a) whether wh-adjuncts are associated with online storage costs similarly to wh-arguments; b) whether there are behavioral differences in processing why compared to other wh-adjuncts, and c) at which point a wh-adjunct gap is actually filled, given that, unlike wh-arguments always associated with their selecting predicate, wh-adjuncts are typically associated with abstract syntactic nodes not directly recoverable from the stimulus.

The study: We conducted two self-paced reading experiments investigating processing of wh-dependencies headed by wh-adjuncts in embedded wh-questions. In the Slovenian Experiment 1, we concentrated on when, and we compared clauses introduced by one of the four main verbs like know which can take either a +wh or –wh clausal complement, viz. The boss knew [comp that / when] the journalist ordered the drink of wine in the bar (Slovenian was chosen because it lexically distinguishes between when as a truly interrogative wh-phrase (“kdaj”) and when as a non-wh operator (“ko”) as in John left when we came, irrelevant here). We expected the when condition to elicit greater online storage costs than the that condition. In the similarly designed English Experiment 2, we manipulated the type of Comp across the values that, how_quickly and why, resulting in 3 conditions (cf. The boss knew [comp that / how_quickly / why] the journalist ordered the drink of wine in the bar), while controlling for the number of words, their frequency and length. We expected that subjects will read the region between the wh-filler and the anticipated gap slower 1) in the how_quickly vs. that conditions; 2) in the how_quickly vs. why conditions, but 3) not in the why vs. that conditions, reflecting the relevant storage costs. Native adult subjects (Exp.1: N= 69, excluded 15; Exp.2: N= 86, excluded 9) read 24 target sentences interspersed with fillers (Exp.1: 62; Exp.2: 50) of similar length, each sentence presented word by word in the center of the screen (to avoid potential distance cues marking a wh-dependency). One third to half of the fillers were followed by comprehension questions.

Results: In both experiments, an extended slowdown in response time was observed in the region R between matrix Comp and the embedded object. Pairwise ANOVAs on cumulative residual reading times showed, in Experiment 1, that R was read slower in the when condition than in the that condition (F= 7.523, p=0.008**). In experiment 2, 1) R was read slower in the how_quickly condition than in the that condition (F=5.021, p=0.029*); 2) a similar slowdown was observed in the how_quickly condition vs. the why condition (F=5.916, p=0.019*) 3) there was no significant difference between the why and that conditions (F=0.00, NS). These diverging results would seem puzzling at face value, but receive a natural explanation in the context of the recent advances in grammatical theory, suggesting that online storage costs are a good indicator of a long-distance dependency even when positing a gap requires online computation of an abstract syntactic node, as is the case with wh-adjuncts.
SARCASM COMPREHENSION: THE EFFECTS OF LITERALITY, FAMILIARITY, AND EXPECTATION

Alexandra Țurcan (University of Nottingham) & Ruth Filik (University of Nottingham)
lpxat3@nottingham.ac.uk

Sarcasm is commonly used in our everyday interactions. However, currently little research has investigated the on-line cognitive processes underlying sarcasm comprehension. We report two eye-tracking experiments, testing the predictions of three theories: the Standard Pragmatic Model (SPM; Grice, 1975), which predicts that sarcastic comments should always take longer to process than literal ones, the Graded Salience Hypothesis (GSH; Giora, 1997), which predicts that sarcastic comments should only take longer to process than literal ones if they are unfamiliar, and the Implicit Display Theory (IDT; Utsumi, 2000), which predicts that an unmet explicit expectation in the context should make sarcastic comments as easy to process as literal ones. Factors manipulated in Experiment 1 were (1) whether the final comment was intended literally, or was an unfamiliar sarcastic comment, and (2) whether comments appeared in a context containing an explicit expectation known to both the speaker and the recipient, or in which the expectation was implicit. Pre-testing the materials insured the effectiveness of our manipulations. Here is one example material:

**Literal comment with explicit expectation:** Dean and Chloe were on holiday in Valencia. The end of the trip was approaching so Dean asked Chloe to think of something thrilling to do on their last day. Chloe suggested they go and watch the Formula 1 race, which was Dean’s favourite sport. “Your suggestion is stirring!” Dean said to her.

**Literal comment with implicit expectation:** Dean and Chloe were on holiday in Valencia. Their trip was quickly coming to an end, and they weren’t sure what to do on their final day. Chloe suggested they go and watch the Formula 1 race, which was Dean’s favourite sport. “Your suggestion is stirring!” Dean said to her.

**Sarcastic comment with explicit expectation:** Dean and Chloe were on holiday in Valencia. The end of the trip was approaching so Dean asked Chloe to think of something thrilling to do on their last day. Chloe suggested they stay in the hotel and watch TV, which was quite boring. “Your suggestion is stirring!” Dean said to her.

**Sarcastic comment with implicit expectation:** Dean and Chloe were on holiday in Valencia. Their trip was quickly coming to an end, and they weren’t sure what to do on their final day. Chloe suggested they stay in the hotel and watch TV, which was quite boring. “Your suggestion is stirring!” Dean said to her.

The critical region was the word that disambiguated each comment as either sarcastic or literal (e.g. “stirring”). Results showed shorter first-pass and regression path reading times for literal than sarcastic conditions, regardless of whether an explicit expectation was present in the context, which supports the predictions of the SPM and the GSH, but not the IDT. In Experiment 2, familiarity was included as an additional factor in the design, in order to further distinguish the predictions of the GSH and the SPM. In the familiar conditions, the final comment of the example material above was “So excited!” which pre-tests indicated to be a familiar irony. The results supported the GSH: in the early stages of processing (regression path reading times of the disambiguating word), familiar sarcastic comments were read as quickly as literal comments, while unfamiliar sarcastic comments took longer to read. There was again no effect of expectation. Interestingly, the processing advantage for familiar sarcastic comments was not observed in the later stages of processing (total reading times of the disambiguating word, and reading times on the post-critical region, e.g. “Dean said to her”), when sarcastic comments were more difficult to process than literal ones irrespective of their familiarity. These results potentially indicate that in the later stages of sarcasm comprehension, there may be an on-going conflict between the literal and ironic meanings of the comment. The GSH best accommodates these results.
PREDICTION USING PRODUCTION OR PRODUCTION ENGAGING PREDICTION?

Florian Hintz, Antje S. Meyer & Falk Huettig (Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands)
florian.hintz@mpi.nl

Prominent theories of predictive language processing assume that language production processes are used to anticipate upcoming linguistic input during comprehension (Dell & Chang, 2014; Pickering & Garrod, 2013). Here, we explore the converse case: Does a task set including production in addition to comprehension encourage prediction, compared to a task only including comprehension? To test this hypothesis, we conducted a cross-modal naming experiment (Experiment 1) including an object naming task and a self-paced reading experiment (Experiment 2) that did not include overt production. We used the same predictable (N = 40) and non-predictable (N = 40) sentences in both experiments. The sentences consisted of a fixed agent, a transitive verb and a predictable or non-predictable target word (The man drinks a beer vs. The man buys a beer). Most of the empirical work on prediction used sentences in which the target words were highly predictable (often with a mean cloze probability > .8) and thus it is little surprising that participants engaged in predictive language processing very easily. In the current sentences, the mean cloze probability in the predictable sentences was .39 (ranging from .06 to .8; zero in the non-predictable sentences). If comprehenders are more likely to engage in predictive processing when the task set involves production, we should observe more pronounced effects of prediction in Experiment 1 as compared to Experiment 2. If production does not enhance prediction, we should observe similar effects of prediction in both experiments.

In Experiment 1, participants (N = 54) listened to recordings of the sentences which ended right before the spoken target word. Coinciding with the end of the playback, a picture of the target word was shown which the participants were asked to name as fast as possible. Analyses of their naming latencies revealed a statistically significant naming advantage of 106 ms on predictable over non-predictable trials. Moreover, we found that the objects’ naming advantage was predicted by the target words’ cloze probability in the sentences (r = .411, p = .016). In Experiment 2, the same sentences were used in a self-paced reading experiment. To allow for testing of potential spill-over effects, we added a neutral prepositional phrase (buys a beer from the bar keeper/drinks a beer from the shop) to each sentence. Participants (N = 54) read the sentences word-by-word, advancing by pushing the space bar. On 30% of the trials, comprehension questions were used to keep up participants’ focus on comprehending the sentences. Analyses of participants’ target and post-target reading times revealed numerical advantages of 6 ms and 20 ms, respectively, in the predictable as compared to the non-predictable condition. However, in both cases, this difference was not statistically reliable (t = .757, t = 1.43) and the significant positive correlation between an item’s naming advantage and its cloze probability as seen in Experiment 1 was absent (r = .037, p = .822). Importantly, the analysis of participants’ responses to the comprehension questions, showed that they understood the sentences (mean accuracy = 93%).

To conclude, although both experiments used the same sentences, we observed effects of prediction only when the task included production. In Experiment 2, no evidence for anticipation was found although participants clearly understood the sentences and the method has previously been shown to be sensitive to measure prediction effects (Van Berkum et al., 2005). Our results fit with a recent study by Gollan et al. (2011) who found only a small processing advantage of predictive over non-predictive sentences in reading (using highly predictable sentences with a cloze probability > .87) but a strong prediction effect when participants read the same sentences and carried out an additional object naming task (see also Griffin & Bock, 1998). Taken together, the studies suggest that the comprehenders’ task set exerts a powerful influence on the likelihood and magnitude of predictive language processing. When the task set involves language production, as is often the case in natural conversation, comprehenders might engage in prediction to a stronger degree than in pure comprehension tasks. Being able to predict words another person is about to say might optimize the comprehension process and enable smooth turn-taking.
MANDARIN SUBJECT RELATIVES ARE EASIER TO PROCESS THAN OBJECT RELATIVES EVEN IN DISCOURSE CONTEXTS
Fuyun Wu (Shanghai International Studies University, SISU) & Yanan Sheng (SISU) 
fywu@shisu.edu.cn

Research in the past decade on prenominal relative clauses (RCs) in Chinese has yet to give a definite answer to the question of whether subject RCs are easier to process than object RCs. Some research suggests that subject RCs are easier ([7, 9]), supporting experience-based theories ([4, 6, 8]). But there is also other research – best represented by two self-paced reading studies ([3, 5]) – showing a processing advantage of object RCs, as predicted by Gibson’s Dependency Locality Theory (DLT, [2, 5]). However, these two studies (as well as others) used bare nominals in RCs, potentially violating presuppositions underlying RC usage for identifying one identifiable referent out of a set, especially in the presence of RC-biasing contexts ([3, 9]). Furthermore, a (distal) demonstrative-classifier phrase in Chinese naturally fulfills a definiteness-denoting function, and is naturally expected in the presence of an RC ([11]).

To address this concern and to further clarify the issue of subject-object processing asymmetry, we conducted a self-paced reading experiment (N=67) with discourse contexts (1) that would presumably bias participants towards an RC parse. We critically manipulated definiteness-denoting dem-classifier sequence (presence vs. absence), in addition to the classic manipulation of extraction types (subject vs. object relatives), yielding (2). Twenty target items adapted from [3, 9] were intermixed with 32 filler items by a Latin Square design.

Following [3, 9], we focus on the RC and head noun as the critical regions. At the first word within the RC (i.e., RC-internal N/V), we found a significant main effect of definiteness (t = 5.49, pMCMC < .0001), an interaction between definiteness and RC type (t = -2.85, pMCMC = .002), but no effect of RC type. Bare object relatives were processed slowest, suggesting that the lack of a definite phrase incurred a higher surprisal in object RCs than in subject RCs. In addition, combined regions of the first two words (NV/VN) yield a marginal main effect of definiteness (t = 1.79, pMCMC = .074) and no other effects, showing that definite RCs were processed faster than bare RCs. Critically at the head noun and at the following region ‘is’, we found a significant main effect of RC type (ts > 2.2, ps < .05) and no other effects (ts < 1): Subject relatives were processed faster than object relatives. This finding is crucially different from that reported in [3, 9], and is contrary to the prediction made by the DLT’s retrieval (integration)-cost metric.

Thus, our finding shows the importance of having the definiteness-denoting phrase in pronominal RCs in Chinese, to not only license the discourse function of RCs, but also help resolve the debate on subject-object processing asymmetry. Our results are in line with earlier findings ([7, 9, 10, 11]), supporting experience-based approaches ([4, 6, 8]).

(1) Two young men went to a bar for a drink after a friend’s birthday party./Because of some verbal conflicts, a waiter punched one young man, and then the other young man punched the waiter./ Xiaoming said: I heard that the bar owner knew one of the young men and the bartender knew the other. Which young man was the one that the bar owner knew?!!
(2) a/c: bare/definite object relatives
   Xiaomei says: (that-CL)|waiter|hit-ASP|DE|young man|is|boss|knew|SFP.
   b/d: bare/definite subject relative:
   Xiaomei says: (that-CL)|hit-ASP|waiter|DE|young man|is|boss|knew|SFP.

Disfluencies (such as *uh* and *uhm*) are a common phenomenon in spontaneous speech. Rather than filtering these hesitations from the incoming speech signal, listeners are sensitive to disfluency and have been shown to actually use disfluencies for speech comprehension. For instance, disfluencies have been found to have beneficial effects on listeners’ memory [1]. Accumulating evidence indicates that attentional mechanisms underlie this disfluency effect: upon encountering disfluency, listeners raise their attention to the incoming speech signal [2].

The experiments reported here investigated whether these beneficial effects of disfluency also hold when listening to a non-native speaker. Recent studies on the perception of non-native disfluency suggest that disfluency effects on prediction are attenuated when listening to a non-native speaker [3]. This attenuation may be a result of listeners being familiar with the frequent and more variant incidence of disfluencies in non-native speech. If listeners also modulate the beneficial effect of disfluency on memory when listening to a non-native speaker, it would indicate a certain amount of control on the part of the listener over how disfluencies affect attention, and thus comprehension. Furthermore, it would argue against the hypothesis that disfluencies affect comprehension in a rather automatic fashion (cf. the Temporal Delay Hypothesis; [4]).

Using the Change Detection Paradigm, we presented participants with three-sentence passages that sometimes contained a filled pause (e.g., “... that the patient with the *uh* wound was...”). After each passage, participants saw a transcript of the spoken passage in which one word had been substituted (e.g., “wound” > “injury”). In our first experiment, participants were more accurate in recalling words from previously heard speech (i.e., detecting the change) if these words had been preceded by a disfluency (relative to a fluent passage). Our second experiment - using non-native speech materials - demonstrated that non-native *uh’s* elicited an effect of the same magnitude and in the same direction: when new participants listened to a non-native speaker producing the same passages, they were also more accurate on disfluent (as compared to fluent) trials.

These data suggest that, upon encountering a disfluency, listeners raise their attention levels irrespective of the (non-)native identity of the speaker. Whereas listeners have been found to modulate prediction effects of disfluencies when listening to non-native speech, no such modulation was found for memory effects of disfluencies in the present data, thus potentially constraining the role of listener control in disfluency processing. The current study emphasizes the central role of attention in an account of disfluency processing.

**References:**

TIP-OF-THE-TONGUE IN A SECOND LANGUAGE: 
THE EFFECTS OF BRIEF L1 EXPOSURE ON L2 PRODUCTION 
Hamutal Kreiner (Ruppin Academic Center) & Tamar Degani (Haifa University) 
hamutalk@ruppin.ac.il

One of the most frustrating difficulties bilinguals experience during production is the Tip-of-the-Tongue (TOT) state (Brown & McNeill, 1966), which is a temporary difficulty in retrieving words in one of the languages they speak. Interestingly, TOT incidents are not confined to bilinguals, but bilinguals consistently exhibit higher TOT rates than monolinguals (e.g., Gollan & Acenas, 2004). The research that investigated bilingual TOT has initially focused on general long-term differences in language experience between bilingual and monolingual speakers (e.g., Gollan & Acenas, 2004). However, many bilinguals shift flexibly from one language to another, and such brief language exposure may similarly affect bilinguals' TOT rates.

In the current study, we examine how TOT rates are modulated by both long term and very recent short-term language exposure. Further, we test to what extent global, non-item-specific, effects may result from such brief language exposure. Russian-Hebrew bilinguals who acquired Hebrew either early (<5 years, n=24) or late (>11 years, n=24) were compared to native Hebrew Speakers (n=24) on a picture naming task in Hebrew. TOT rates were examined in two blocks, before and after viewing a short 10-minute movie in Russian.

Both short-term exposure and long-term language experience modulated TOT rates. Before the movie, late bilinguals exhibited significantly higher TOT rates than native Hebrew Speakers and early bilinguals. Early bilinguals, however, did not significantly differ from native Hebrew Speakers. Critically, following the Russian movie, bilinguals in both groups differed from the native Hebrew Speakers. Thus, exposure to the non-target language, which did not involve participants' production or exposure to the particular items to be named in the following task, exerted a global, non-item-specific, cross-language interference effect. The findings highlight the dynamic nature of the bilingual system in which both transient and sustained language factors operate to influence bilingual performance.

References


A semantic relatedness task was employed in two experiments to explore how listeners reinterpret their understanding of sentences in light of new information. Listeners heard sentences containing an ambiguous word (e.g., "toast" that refers to grilled bread or a celebratory speech). To induce reinterpretation, the disambiguating context was presented several words after the ambiguous word and corresponded to the less frequent (subordinate) meaning of the ambiguous word (e.g., "the woman had to make the toast with a very old microphone"). Psycholinguistic research converges on the finding that listeners rapidly select the most frequent meaning of an ambiguous word (e.g., grilled bread meaning) in the absence of prior disambiguating information and, thus, listeners will need to reinterpret their understanding of the sentence when they encounter subordinate-related information ("microphone"). After the sentence, participants were presented with a visual word probe (Experiment 1) or a picture probe (Experiment 2). Participants' task was to decide whether the probe was related or unrelated to the meaning of the sentence. The critical probes were related to either the appropriate meaning of the ambiguous word ("speech") or to the inappropriate meaning ("bread"). Experiment 1 showed that when the probe was presented 100 ms post sentence offset responses were 1) significantly slower for probes that followed ambiguous sentences than unambiguous sentences and 2) additionally slower for contextually-inappropriate probes. Interestingly, the former, general processing cost of ambiguity was significantly reduced when the probe was presented 200 ms later, at 300 ms post-sentence offset, but the latter, more specific interference cost from the inappropriate meaning was only reduced when the probe was presented 1000 ms post-sentence offset. These results suggest that reinterpreting the meaning of sentences is a time-consuming process but that it quickly progresses within 200 ms of hearing the disambiguating word. However, contextually-inappropriate meanings take much longer to be suppressed. Experiment 2 replicated this dissociation between the activation of the contextually-appropriate meaning and inappropriate meanings by the finding of a significant RT cost for contextually-inappropriate picture probes in the absence of an ambiguity cost for contextually-appropriate picture probes. These dissociations are inconsistent with models that suggest a very tight coupling between the activation of the alternative meanings of ambiguous words, and hence, suggest a degree of independence between their activation levels during sentence comprehension.
AGREEMENT ATTRACTION IN MODERN STANDARD ARABIC COMPREHENSION
Matthew A. Tucker (NYU Abu Dhabi), Ali Idrissi (United Arab Emirates University), and Diogo Almeida (NYU Abu Dhabi)
matt.tucker@nyu.edu

Data from errors in the processing of verb-argument agreement dependencies are crucial to theorizing about language processing (see [1,2] and references therein). However, all such studies have examined these errors in languages of the Indo-European family. The present work investigates the appearance of such errors in comprehension in Modern Standard Arabic grammatical number morphology via self-paced reading. We demonstrate that the kind of plural formation on NPs in agreement attraction configurations (suffixing vs. internal vowel change) can modulate attraction effects.

Participants read 48 sentences which contained a subject relative clause modifying the subject. The morphological number of the object of the relative clause (the distractor NP) and the main clause verb was varied systematically for singular and plural. Stimuli were counterbalanced for NP gender (masculine, feminine) and included an adverb between the distractor NP and main clause verb ([2]). All subjects were singular. For example:

ʔal-mutarʒim-u ʔallaðii saaʕad-a ʔal-raʔiis-a ʔaħjaanan
the-translator-nom comp.masc.sg helped-3.sg.masc the-president-acc often
ja-takallamu xamsata luɣaat-in bi-faʕaahatin.
3.sg.masc-speaks five languages-acc with-fluency

“The translator who helped the president often speaks five languages fluently.”

The design yields four conditions: 1) SS (sg distractor, sg verb), 2) SP (sg distractor, pl verb), 3) PS (pl distractor, sg verb), and 4) PP (pl distractor, pl verb).

Analysis reveals predicted agreement attraction effects. In the critical verb region, both ungrammatical conditions (SP and PP) were read more slowly than grammatical conditions (SS and PS; $F(1,82) = 45.65, p < 0.001$). However, we also observe an expected interaction: the presence of a plural distractor in the PP cases leads to lower mean reading times relative to fully ungrammatical baselines in the SP condition. This effect is significant in the critical verb region ($F(1,82) = 4.015, p < 0.05$) and marginal in the first spillover region (see the Verb and Verb + 1 regions in the figure below). We observe that this attraction effect is driven entirely by feminine plurals, which are all suffixal, whereas masculine plurals, which are formed by internal vowel change in our stimuli, display no significant attraction effect (not shown in plot below).

Finally, we observed a clear processing cost of plural nouns reflected in significantly increased reading times for PS and PP conditions in regions preceding the target verb, in line with [2] (see the Distractor and Adverb regions, below; $F(1,82) = 8.66, p < 0.01$ for distractor region). Similar results are obtained using residualized reading times as with raw reading times.

These results support the findings of [2] that plural NPs involve a processing cost as well as the universality of agreement attraction effects. However, differences in the behavior of masculine versus feminine plurals support a language-specific notion of morphological plurality such that attraction effects can be sensitive to the type of plural formation seen on NPs.

One of the fundamental processes in language comprehension is the processing of dependencies between grammatical elements such as agreement relations between verbs and their arguments. Understanding the time course and localization of this processing is a central goal of neurobiological investigations of language comprehension. To this end we conducted an experiment which investigated the visual processing of subject-verb agreement dependencies in English with concurrent EEG and MEG recording. We selected stimuli which included an agreement attraction configuration wherein a “distractor” non-subject noun provided the possibility for erroneous grammatical parses assigned to ungrammatical verbs. Virtually no neurolinguistic studies examine agreement dependencies, let alone attraction configurations (especially in the MEG methodology).

19 (of 24 planned) participants read 384 sentences with a prepositional modifier attached to a subject (The key(s) to the cabinet(s) obviously was/were rusty from years of disuse.) in a rapid serial visual presentation paradigm. The grammatical number of the subject noun (key(s)), non-subject distractor noun (cabinet(s)), and target verb (was/were) were fully crossed in a 2x2x2 design. After each sentence, subjects performed an acceptability judgment. The acceptability responses reveal a clear attraction result: compared to wholly unacceptable sentences, acceptability ratings are higher when the two nouns are mismatched, and plural distractor nouns with singular subjects yield higher error rates than singular distractors with plural subjects.

ERP analysis reveals a positive-going deflection in ungrammatical configurations relative to grammatical configurations in the 500-700 ms latency range. This positivity is centro-posterior in distribution and is therefore interpreted as an instance of the P600 effect typically observed for ungrammatical agreement dependencies. Moreover, this P600 is reduced in amplitude under the same conditions driving higher acceptability ratings: the P600 for plural distractors is reduced relative to singular distractors and fully ungrammatical sentences where the verb matches neither noun (the attraction configuration).

MEG analysis reveals a decrease in the evoked magnetic response to ungrammatical sentences relative to grammatical ones at the same latency as the observed electrical P600. This decreased activation is higher at occipital sites and is driven by an attraction-configuration change overlaid on top of the M170 to the word immediately following the target region. Additionally, similar patterns of activation are found at the target verb itself in the pars triangularis (BA 45) and superior temporal gyrus (BA 22p/41/42).

The results of this study argue for a parsing-mediated interpretation of the P600 given the reduction in amplitude that correlates with behavioral errors indicative of the illusion of ungrammaticality. Furthermore, the MEG results reveal that there are ERF correlates of grammatical dependency violations which can be observed at both occipital and left-temporal locations in the 500-700ms post-stimulus window. Finally, the results here suggest that MEG activation is inversely correlated with the success of dependency processing (and therefore the ERP response), a fact which is consistent with the emerging literature on MEG correlates of grammatical dependency processing.
A core aspect of speech production is how speakers combine individual words to multi-word phrases. Yet little is known about the mechanism underlying phrase production. In the present study, we focused on noun phrase production (i.e., adjective + noun; red box). I examined whether speakers plan and retrieve noun phrases as a whole before speaking (Global Production Hypothesis) or they incrementally produce each word (Incremental Production Hypothesis).

To test these two hypotheses, twenty-seven native speakers of Dutch were asked to name 461 colored objects (10 colors combined with 50 objects). I measured phrase onset (i.e., the time from the picture onset to the moment participants started speaking), eye-gaze duration (i.e., eye-gaze duration on the colored object), and the time lag between the adjective and the noun onset. Multiple regression analyses was carried out to examine the impact of the phrase frequency, the concept familiarity of the phrase, the phrase length, and the adjective frequency, and the noun frequency. Table 1 summarizes results.

Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Phrase onset</th>
<th>Eye-gaze duration</th>
<th>Time lag between words' onsets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phrase frequency</strong></td>
<td>β = -.12</td>
<td>-.14</td>
<td>-.17</td>
</tr>
<tr>
<td>t-value</td>
<td>-2.45*</td>
<td>-2.93**</td>
<td>-3.59**</td>
</tr>
<tr>
<td><strong>Phrase length</strong></td>
<td>β = -.01</td>
<td>.23</td>
<td>.37</td>
</tr>
<tr>
<td>t-value</td>
<td>-.24</td>
<td>4.58**</td>
<td>7.78**</td>
</tr>
<tr>
<td><strong>Phrase concept familiarity</strong></td>
<td>β = -.34</td>
<td>.00</td>
<td>.33</td>
</tr>
<tr>
<td>t-value</td>
<td>-3.05**</td>
<td>.04</td>
<td>7.67**</td>
</tr>
<tr>
<td><strong>Adjective frequency</strong></td>
<td>β = -.01</td>
<td>-.20</td>
<td>.11</td>
</tr>
<tr>
<td>t-value</td>
<td>-7.36**</td>
<td>4.22**</td>
<td>2.55*</td>
</tr>
<tr>
<td><strong>Noun frequency</strong></td>
<td>β = -.13</td>
<td>-.00</td>
<td>.19</td>
</tr>
<tr>
<td>t-value</td>
<td>-.22</td>
<td>-.06</td>
<td>3.77**</td>
</tr>
<tr>
<td><strong>R square</strong></td>
<td>.19</td>
<td>.16</td>
<td>.22</td>
</tr>
<tr>
<td><strong>F-value</strong></td>
<td>21.8**</td>
<td>17.26**</td>
<td>25.51**</td>
</tr>
</tbody>
</table>

Note: **p < .01, *p < .05.**

Phrase onset and eye-gaze duration were predicted by both the phrase and the adjective frequency. Furthermore, the time lag between words onset was predicted by the noun frequency as well as other predictors. These results suggest that speakers plan and retrieve both the adjective and the whole phrase before speaking and speakers only retrieve the noun after phrase onset. Thus, neither the Global Production Hypothesis nor the Incremental Production Hypothesis can fully explain the current results. To conclude, noun phrase production is probably a mixture of global and incremental processes.
DEVELOPMENT IN PERSPECTIVE-TAKING DURING AMBIGUTIY AVOIDANCE
Kumiko Fukumura, Lauren Ferguson, Katie Quinn (University of Strathclyde)
kumiko.fukumura@strath.ac.uk

Children aged around 5-to-6 years can produce contrastive modifiers by taking their addressee’s visual perspective into account; they produce size adjectives (as in “small door”) more when the addressee can also see another category exemplar that contrasts in size (a large door) than not [1]. But we know little about development after children start using the addressee’s perspective when producing complex noun phrases.

To investigate how children become adult-like communicators, Exp 1 contrasted 6-10 years old children with and without Autistic Spectrum Disorder (ASD) and adults without ASD. Children with ASD are said to be selectively impaired with perspective-taking or mentalizing compared to children without ASD [2], and children without ASD also have more difficulty with overcoming their own, “egocentric” perspective than adults [3]. Indeed, children with ASD produced size adjectives (“small door”) significantly more than the age/ability-matched children without ASD when the size-contrasting competitor (large door) was hidden from their addressee (privileged context). But interestingly, both groups of children produced disambiguating size adjectives equally often when the competitor was also visible to and hence shared with their addressee (see Fig 1). In the same task, adults almost always avoided referential ambiguity in the shared context and almost never produced redundant size adjectives in the privileged context. Exp 2 compared adolescents with ASD, aged 11-16, with their peers without ASD. As in Exp 1, adolescents with ASD produced redundant size adjectives significantly more than adolescents without ASD in the privileged context, but they did not significantly differ in the shared context. Non-ASD adolescents improved significantly on ambiguity avoidance in the shared context compared to younger children, and they differed only marginally from adults. But overall, ASD and Non-ASD adolescents produced adjectives as often as their younger counterparts in the privileged ground condition.

In sum, the ASD group produced size adjectives as often as the non-ASD group in the shared context, suggesting that autism does not impair ambiguity avoidance itself [contra 4]: When the addressee could see what speakers could see, whether speakers adopt the addressee’s perspective does not influence the production of felicitous reference, and contrary to traditional but still dominant views, unique reference may not require mentalizing abilities. In contrast, avoiding size adjectives in the privileged context requires perspective-taking: the use of size adjectives was appropriate to the speaker’s view but not to their addressee’s. Therefore, children with ASD produced egocentric descriptions more than children without ASD, and adolescents without ASD also improved less on avoiding egocentric descriptions than on avoiding ambiguous descriptions. Additionally, although children/adolescents occasionally corrected ambiguous bare nouns (“door, no, large door”), they almost never corrected the contrastive adjectives when their addressee could not see the contrast (“large door, door”), indicating that children were less likely to realize the infelicity of their egocentric descriptions (or the oddity of “small” when the referent does not look that small to their addressee). To conclude, adult-like addressee adaptations take many years to master: avoiding egocentric descriptions requires perspective-taking that challenges children, particularly those with ASD, delaying its learning relative to unique reference.

SYNTACTIC PRIMING: A LEXICAL BOOST, CUMULATIVITY, AN INVERSE PREFERENCE EFFECT AND... A POSITIVE PREFERENCE EFFECT
Katrien Segaert (MPI for Psycholinguistics) and Peter Hagoort (MPI for Psycholinguistics) katrien.segaert@mpi.nl

Introduction. Syntactic choices are sensitive to priming. The following factors have been repeatedly identified to modulate priming of choices: the lexical boost (lexical repetition boosts priming), cumulativity (multiple primes increase the effect) and the inverse structure preference effect (less preferred structures prime more). Priming does not only influence syntactic choices however, it also affects production latencies. Interestingly, studies investigating syntactic priming of latencies recently demonstrated a positive structure preference effect: preferred structures prime more\(^1\)\(^2\). The positive preference effect for latencies is stable across experiments and is demonstrated for transitives\(^1\) and ditransitives\(^2\).

In two experiments we examined whether syntactic priming of latencies is sensitive to the same factors known to influence priming of choices (the lexical boost, cumulativity and structure preference). If all these phenomena indeed influence priming of both choices and latencies, it supports that latencies are a valid syntactic priming measure and that syntactic priming theories should include mechanisms to explain latency effects.

Method and results. We investigated active/passive voice alternation using a picture description paradigm simultaneously measuring production choices and latencies. We analyzed the data using logit and linear mixed models.

Experiment 1 (N=45) investigated if syntactic priming is affected by the cumulative effect of the number of immediate primes (1 vs. 3 primes) and the cumulative effect of all preceding target productions within the experiment (i.e. the proportion of passives out of total active and passive response productions in the experiment so far, hereafter: CumPassProp). In the baseline measure, fewer passives (8.2%) than actives were produced (p<.001). Response choices were not affected by active primes (1 prime: p>.3, 3 primes: p>.5) but were affected by passive primes (1 prime: p<.001, 3 primes: p>.001). More passive productions followed 3 passive primes than 1 (p<.006). Also, the higher CumPassProp, the more passives produced (p<.004). Response latencies for active structure choices were predicted by syntactic repetition (p<.003) and this effect marginally interacted with CumPassProp (p<.08): the higher CumPassProp (i.e. the smaller the cumulative proportion of actives), the smaller the priming effect for actives. Latencies for passives were unaffected by priming (p>.3).

Experiment 2 investigated if priming is modulated by verb repetition and by CumPassProp (N=45). In the baseline measure, fewer passives (7.8%) than actives were produced (p<.001). Response choices were not affected by active primes (novel verb: p>.4, repeated verb: p>.1) but were affected by passive primes (novel verb: p<.001, repeated verb: p<.001) and by CumPassProp (p<.001). Also, more passives were produced following a passive prime with the same verb than a different verb (p<.001). Response latencies for active structure choices revealed a syntactic priming effect (p<.001). For passives there was only latency priming when aided by verb repetition and a high CumPassProp value (3-way interaction syntactic repetition x verb repetition x CumPassProp: p<.04).

Conclusion. Syntactic priming in choices and latencies is modulated by the same factors: the lexical boost, cumulativity and structure preference. However, while structure choices show inverse structure preference effects, latencies show positive preference effects. Well-established syntactic priming accounts with an implicit learning component\(^3\)\(^4\) focus on syntactic choices and explain cumulativity and the inverse preference effect, some of these also explain the lexical boost. This cannot be the full story however: latencies reveal that preferred structures benefit from priming as well. A mechanism able to explain this -a competition\(^1\)\(^2\) or other mechanism with the same explanatory power- could be incorporated in existing accounts of syntactic priming\(^3\)\(^5\) which so far have only focused on choices.

EFFECTS OF NEIGHBOURHOOD DENSITY ON LEXICALITY JUDGEMENTS ACROSS THE LIFESPAN AND IN INDIVIDUALS WITH ALZHEIMER’S DISEASE

Nancy Azevedo (McGill University; CRIR-JRH), Eva Kehayia (McGill University; CRIR-JRH), Ruth Ann Atchley (University of Kansas) & Vasavan N.P. Nair (McGill University; Douglas Mental Health University Institute)
nancy.azevedo@mail.mcgill.ca

Introduction: Neighbourhood Density (N), i.e. the number of lexical neighbours a word has, is a variable that impacts visual lexical processing in English. As one progresses through the lifespan, qualitative changes in the processing of lexicality occur (Ratcliff et al., 2004), as seen in differences in how N is processed by older adult readers (Stadtlander, 1995). In young adults, the effect of N on lexicality judgements is paradoxical: a large N facilitates real word access but is inhibitory for pseudowords. In contrast, results with older adults report faster response times (RTs) to real words with a low neighbourhood density or no difference in RT to pseudowords based on N (Stadtlander, 1995). However, there is little research investigating N effects in healthy aging. Further, to our knowledge, there are no studies examining this effect in individuals with Alzheimer’s disease (AD), though on-line lexical processing studies do report slower RTs in those with AD (Taler & Phillips, 2008). To help fill this gap in the literature, we investigated lexicality judgements in English-speaking young adults (YA; n=35), older adults (OA; n=19), and individuals with AD (n=15) focusing on neighbourhood density.

Methods: Participants performed a 320-trial lexical decision task, during a single session, probing three stimuli types: words (80 nouns, PENCIL and 80 verbs, CANCEL), pseudowords (80 legal/pronounceable letter strings, GLONET), and nonwords (80 illegal/unpronounceable letter strings, TLKMP). Stimuli were controlled for frequency, length and number of lexical neighbours. Experimental word (nouns) and pseudoword stimuli were divided into high and low N groups (40/group). High density stimuli had 5-18 neighbours, low density stimuli had 0-2, and nonwords had 0. RTs and accuracy (AC) were recorded for each stimulus.

Results and Conclusions: Overall, we observed a progressive slowing of reaction times (RT) across all stimulus groups that occurs with age and becomes more prominent with AD. Interestingly, this slowing may help improve accuracy in the aging groups, particularly for the more challenging pseudowords. With regard to N, results for the young adults mirror what has previously been reported in the literature. Also, all three groups were similarly sensitive to N for pseudowords in both AC and RT. However, older individuals in our study are less sensitive to N when they were making a lexical decision for a real word, and this is further altered by the presence of AD. A large N was less facilitatory for words for older adults’ reaction times and there was no accuracy advantage. For individuals with AD there was no facilitation (in AC or RT) for high N words. We propose that a slowing of activation mechanisms that occurs with age and is aggravated by the AD degenerative process may be contributing to the dissociation of N effects in words versus pseudowords, observed in the current study.


HOW LANGUAGE MEDIATES THE CREATION OF FALSE MEMORIES

Agnieszka E. Konopka (Max Planck Institute for Psycholinguistics)
agnieszka.konopka@mpi.nl

How does linguistic proficiency influence memory for verbal information? Sampaio and Konopka (2012, Memory) showed that non-native (L2) speakers have better memory for the surface form of simple sentences (verbatim memory) than native (L1) speakers, particularly when sentences include contextually dispreferred words, likely due to more effortful processing of L2 input. The present study tests whether better retention of surface form also has implications for retention of sentence meaning (gist memory) by comparing false memory rates in L1 and L2 speakers.

30 adult L1 speakers and 24 fluent L2 speakers of Dutch (L1 German, mean age of acquisition of Dutch = 19 years) studied pictures presented with recorded descriptions. Target pictures showed transitive events (e.g., a dog chasing a mailman; n=42) and were described with active sentences. The sentences were selected from descriptions produced spontaneously by L1 speakers (n>400) in earlier production experiments, and were either the modal description of these events (preferred [PR] version: The dog ran after the mailman) or a less frequent description (using a synonymous [SYN] verb: The dog chased the mailman). Sentence type (PR vs. SYN) was confirmed by norming. Participants were asked to pay attention to sentence wording for a later memory test.

At test, participants made recognition judgments for a list of studied and unstudied sentences on a 1-5 scale (1=new, 5=old; they were asked to rate sentences with wording changes as new sentences). Test lists included PR sentences, SYN sentences, and novel sentences describing an implied event (unstudied INFERENCEs [INF], e.g., The dog bit the mailman, that differed from PR and SYN items only with respect to the main verb). Recognition memory was compared using d’ in two steps.

The first analysis tested memory for surface form (verbatim memory) by comparing ratings given to PR and SYN sentences. Overall, participants gave higher ratings to studied than unstudied sentences (means = 4.0 and 2.8 respectively). L1 and L2 speakers gave similar ratings to sentences studied in the PR version, both when the sentences were presented in the PR version (hits) and the SYN version at test (false alarms; d’ = 1.2 vs. 1.0 in L1 and L2, p=.4). However, for sentences studied in the SYN version, L2 speakers were more accurate at recognizing the same SYN sentences at test than L1 speakers (d’ = .8 and 1.3 in L1 and L2, p<.05), as expected.

The second analysis assessed memory for sentence meaning (gist memory) by comparing ratings given to studied sentences (hits) and unstudied INF sentences (false alarms) in the two groups. L2 speakers gave lower ratings to INF sentences than L1 speakers (means = 1.9 vs. 2.1). The effect was observed for sentences studied in the SYN version (d’ = 1.6 vs. 2.2 in L1 and L2, p<.05) and not in the PR version (d’ = 1.9 vs. 1.8 in L1 and L2, p=.6). Thus L2 speakers had selectively better memory for sentences originally studied with less preferred wording and were less likely to endorse a corresponding INF sentence as having been studied (a false memory).

The results are the first to show that linguistic proficiency influences the accuracy of memory representations at multiple levels. Fluent L2 speakers outperform L1 speakers in memory for sentence form (verbatim memory) and sentence meaning (gist memory): better retention of surface form details during non-native language processing can reduce the likelihood of committing costly false memory errors (i.e., errors with broader real-world implications than errors in recall of surface form).
FAST AND FURIOUS LEXICAL PROCESSING
Jacolien van Rij (University of Tübingen), Cyrus Shaoul (University of Tübingen), R. Harald Baayen (University of Tübingen), Petar Milin (University of Novi Sad, University of Tübingen)

RSVP experiments (e.g., Rubin & Turano, 1992) suggest that readers can comprehend texts presented with 1000 words/min (17ms p/w). However, the dominant view in morphological processing literature is that reading a complex word is a two-step process, with an early, obligatory and meaning-blind morphemic decomposition (Rastle & Davis, 2008). Also the EEG literature assumes that semantic interpretation begins at approximately 300-500 ms post stimulus onset (but see Segalowitz & Zheng, 2009). To investigate how early in lexical processing readers can use semantic information associated with the word, we performed two experiments using the RSVP paradigm.

Study 1: We tested single word recognition using a masked lexical decision task with varying presentation durations (8, 17, 25, 33, 42, 50, 58, 67 ms). Stimulus words were sampled randomly from an exhaustive list of 1050 German mono- and multi-morphemic words. To cover a broad range of potentially relevant orthographic and semantic measures that were calculated from a corpus of mixed German texts (300 million words). Additionally, we manipulated the type of mask. We used both (a) noise masks, similar to commonly used mask-character; and (b) strings of random consonants.

Data analyses (using Generalized Additive Modeling, n=19) showed that random consonant masks interfered with lexical processing: accuracy only rose above chance at exposition times greater than 60 ms, whereas standard noise masks already yielded above chance words recognition at exposition times at 33 ms. Interestingly, our analyses also revealed significant effects of co-occurrence neighborhood density and of corpus frequency in predicting RTs of successfully recognized words: words that had denser semantic neighborhoods were recognized faster, as were words that had higher corpus frequency. The analysis indicates that semantic effects are present very early in lexical access.

Study 2: We tested recognition of sentence words using a picture/sentence-verification task. Participants (n=60) were asked to judge whether a word illustrated in the picture was in the sentence, which was presented word-by-word, 33 ms per word. The picture followed or preceded the sentence (blocked design). We created 4 variants of each item, manipulating the position of the word by adding an adverbal phrase, and the congruency with the picture. We included a lexical decision task similar to Exp. 1 as pre-test, with only random consonant masks to make the task comparable with the sentence task.

Our analyses show that linguistic and visual context reduce the time it takes to recognize a word: Participants score far above chance when recognizing words in the picture/sentence verification task with 33 ms exposition duration, but they need double the time (more than 60 ms) for recognizing words without context. The high accuracy on the items in which the picture followed the sentence, although lower than in which the picture preceded the sentence, indicates that the context effect is not merely caused by a priming effect from the picture, but that the semantic information is processed.

Thus, we conclude that lexical processing is a “fast and furious” process, where orthographic input triggers immediate semantic interpretation. Moreover, we showed that sentential context (which is generally present during natural reading) reduces processing time even more. These results are not what form-first, meaning-much-later models would lead one to expect, but resonate with form-and-meaning and amorphous processing accounts (e.g. Feldman et al., 2012; Baayen et al., 2011).

References
ANOMALOUS TRANSFER OF SYNTACTIC RULES BETWEEN LANGUAGES
Awel Vaughan-Evans (Bangor University), Jan Kuipers (University of Stirling), Guillaume Thierry (Bangor University) & Manon Jones (Bangor University).
psepe5@bangor.ac.uk

Each human language possesses a set of distinctive syntactic rules. It remains unknown however, whether syntactic processing is performed on the basis of lexical-phonological regularities (e.g. Thierry and Wu, 2007) or due to the application of generalised abstract rules (e.g. Elman, 1998). Here, we tested the possibility of implicit syntactic transfer between languages in early adult bilinguals to determine whether the implementation of abstract linguistic rules relies on lexical-phonological associations or syntactic contingencies.

We recorded electrophysiological brain responses in 19 Welsh-English bilinguals reading English sentences. All 320 test sentences ended in non-words created by substituting the initial consonant of the final word with a consonant that either produced a mutated or an aberrant form, according to the Welsh soft mutation rule (e.g., ‘c’ substituted by ‘g’-mutated, or ‘d’-aberrant). Syntactic structure was manipulated such that the Welsh translation of the English sentences required a word-final mutation or not (see example 1).

Example1.

<table>
<thead>
<tr>
<th>Mutation Context</th>
<th>Correct</th>
<th>Mutated</th>
<th>Aberrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each book starts with a page listing its contents gynnws</td>
<td>gcontents</td>
<td>dontents</td>
<td></td>
</tr>
<tr>
<td>“Dechreuir pob llyfr â thudalen yn rhestru ei gynnws”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Mutation Context</td>
<td>contents</td>
<td>gcontents</td>
<td>dontents</td>
</tr>
<tr>
<td>The lid was lifted to examine the contents gynnws</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Codwyd y caead er mwyn archwilio’r gynnws”</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Manipulating the initial consonants allowed us to distinguish between processing of expected and unexpected initial consonants of the final noun depending on the syntactic context. Thus, we were able to investigate the occurrence of syntactic transfer via the phonological mismatch negativity (PMN; an ERP index sensitive to lexical processing modulated by phonological expectation formed on the basis of a word’s initial letter; Connolly and Phillips, 1994).

We found that English words mutated according to Welsh soft mutation rules are more easily integrated than aberrant words but only within a sentence context which requires a mutation in Welsh. Crucially, this effect was found irrespective of phonological overlap between English words and their Welsh translation equivalents (e.g. ‘contents’ shares an initial consonant with its Welsh translation ‘cynnws’, whereas ‘patients’ does not share an initial consonant with its Welsh translation ‘cleifion’).

These results provide evidence for spontaneous transfer of abstract syntactic representations across languages.

References


A NEUROCOMPUTATIONAL MODEL OF THE N400 AND THE P600

Harm Brouwer (University of Groningen), John Hoeks (University of Groningen)
me@hbrouwer.eu

One decade ago, researchers using event-related brain potential (ERP) measurements stumbled upon what looked like a Semantic Illusion in language comprehension: Semantically anomalous, but otherwise well-formed sentences (1b) did not affect the meaning-related N400 component, but increased the amplitude of the structure-related P600 component instead (that is, relative to control (1a)) [1].

<table>
<thead>
<tr>
<th>Sentence (literal translation from Dutch)</th>
<th>Condition</th>
<th>ERP-effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a) 'The javelin was by the athletes thrown'</td>
<td>Control (passive)</td>
<td>—</td>
</tr>
<tr>
<td>1b) 'The javelin has the athletes thrown'</td>
<td>Reversal (active)</td>
<td>P600</td>
</tr>
<tr>
<td>1c) 'The javelin was by the athletes summarized'</td>
<td>Mismatch (passive)</td>
<td>N400 / P600</td>
</tr>
<tr>
<td>1d) 'The javelin has the athletes summarized'</td>
<td>Mismatch (active)</td>
<td>N400 / P600</td>
</tr>
</tbody>
</table>

This finding spawned five new models of language comprehension, all of which explain the absence of an N400-effect by assuming that the language comprehension system includes an additional semantic processing stream which is not driven by structure, but by word meaning alone. Despite the attractiveness of these multi-stream models, none of them seems capable of explaining the full range of relevant findings in the literature [2]. Most of the models, for instance, have difficulties with semantically anomalous, but syntactically well-formed sentences that produce biphasic N400/P600-effects (sentences (1c) and (1d), as compared to control sentence (1a)).

Brouwer et al. [2,3] put forward a much simpler single-stream model that is able to predict all relevant ERP patterns found in the literature. On their account, the N400 component reflects retrieval of lexical semantic information, rather than compositional semantic processing or semantic integration. Retrieval of information associated with a word, which is presumed to be mediated by the left posterior Middle Temporal Gyrus (lpMTG) [3], is facilitated if that information is already (partly) activated by its prior context. This explains all N400 effects shown in the table above in terms of lexical and contextual priming (thrown is to some extent preactivated by javelin and athletes, but summarized is not). Semantic integration, in turn, is assumed to be mediated by the left Inferior Frontal Gyrus (lIFG), and to be reflected in P600 amplitude. This explains the P600-effects for (1b-d) relative to (1a), as the sentence final verbs in (1b-d) render the sentences semantically anomalous, causing integration difficulty.

We present a neural network model (of the ‘Elman’ type, cf. [4]) that directly instantiates this functional-anatomic mapping of the N400 \( \rightarrow \) lpMTG (lexical retrieval) and P600 \( \rightarrow \) lIFG (semantic integration), and show that this model is able to predict both isolated P600-effects ((1b) versus (1a)), as well as biphasic N400/P600 effects ((1c-d) versus (1a)), providing a proof-of-concept of Brouwer et al.’s Retrieval-Integration account [2,3].

A word's frequency is usually considered to play a central role in visual word recognition. However, recordings of event-related potentials (ERPs) by means of rapid serial visual presentation (RSVP) have revealed a clear impact of frequency on the N400 amplitude only in word lists or at the beginning of single sentences (e.g. Münte et al. 2001), while frequency has no such influence with increasing contextual information, i.e. when words occur later in a sentence. Hence, with unfolding sentence context, the N400 appears to be solely dependent on a word's predictability, as measured via cloze probability (Van Petten & Kutas 1990, 1991). Contrary to this, many studies have found that both frequency and predictability robustly and independently modulate a reader’s eye fixation durations, regardless of a word’s position in a sentence (e.g. Rayner et al. 2004; Kennedy et al. 2013). We aimed to investigate this apparent methodical discrepancy and, in particular, to determine whether the N400 may be sensitive to lexical frequency in context when text is presented normally rather than in RSVP format. The present experiment therefore used a concurrent EEG-eye movement experimental setup to measure fixation-related potentials (FRPs) and fixation durations during reading of the same sentences.

Thirty-eight native speakers of English read 160 single sentences for comprehension while their EEG and eye movements were recorded. Using a 2x2 design, 80 frequent and 80 infrequent target words were each embedded in a unique sentence frame that either did or did not make the critical word predictable. The position of the target averaged approximately ten words into the sentence. Frequencies were based on the SUBTLEX corpus (Brysbaert & New, 2009) and predictability scores were collected with a cloze task. FRPs and fixation durations were examined for the first fixation on the target as well as for the prior fixation. Analyses for the FRPs were restricted to posterior electrodes as these are least affected by eye movement artifacts in the EEG signal (Picton et al. 2000).

Our findings confirm the previously reported lack of an N400 frequency effect, when a word is preceded by substantial sentence context. In fact, frequency did not seem to influence any part of the FRP signal elicited by foveal or parafoveal processing (first fixation and prior fixation). Contrary to this, high predictability from context led to reduced N400 amplitudes for target words and also enhanced the preceding positivity (P200). All predictability effects were triggered by the first fixation on the target. The eye movement results replicated the independent main effects of both factors: the duration of the first fixation on the target was longer for low frequency or unpredictable words, with these effects being strictly additive.

In replicating this well-known discrepancy between methods, we can rule out explanations based on the unnatural reading situation in previous studies using RSVP. The cross-method comparison suggests that N400 and eye movement effects may reflect only partly overlapping aspects of word recognition and that frequency is more important in oculomotor planning than in modality-independent semantic processes in language comprehension, which are assumed to be reflected in the N400 (Kutas & Federmeier 2011).
Integrating a mismatching word elicits characteristic electrophysiological and behavioral responses. Reading studies using eye tracking have shown that the effort to process such words is reflected in longer reading times and regressive eye movements (Clifton, Staub, & Rayner, 2007). In event-related potentials (ERPs), syntactic violations typically elicit a P600 (Osterhout & Holcomb, 1992) and semantic violations an N400 (Kutas & Hillyard, 1980). However, the dominant experimental paradigm in ERP research does not allow participants to adjust their fixation times or to make regressive saccades. As a consequence, it is not clear how these neural and behavioral responses map onto each other. We conducted a study in which we recorded eye movements and brain potentials concurrently while participants read sentences containing a violation. The materials followed the design in Hagoort (2003):

1. Der/The verfallene/neugierige Bauernhof braucht eine Renovierung. Er ist the run-down/inquisitive farm needs a renovation. It is...
2. Der erfahrene Star spielt die/das schwierige/elektrische Rolle. Er überzeugt... The experienced star plays the difficult/electric role. He convinces...

Sentences contained a sentence-medial (1) or sentence-final (2) noun phrase with a gender mismatch between determiner and noun (syntactic violation, violating determiner in bold face, critical noun underlined) or an adjective that did not match with the noun semantically (semantic violation, italics). To obtain a baseline for the natural reading results, we also collected data using RSVP and randomly assigned participants (N=72) to either presentation mode. In both RSVP (N=24) and natural reading (N=48), we replicated Hagoort’s (2003) results. Syntactic violations elicited a P600 in sentence-medial position and an N400-P600 in sentence-final position. Semantic violations engendered an N400-P600 in sentence-medial and an N400 (plus a P600 in natural reading) in sentence-final position. To investigate the neural trace of different eye movement trajectories, we subset the natural reading data into trials with and without first-pass regression from the target noun. Previous work has shown that regressions are closely associated with the P600 (Dimigen et al., 2007). Sentence-medial violations elicited a P600 when a regression occurred but no response without regression. Sentence-final violations yielded an N400-P600 when readers made a regression but a qualitatively different response when they did not: a large sustained negativity at centro-parietal electrodes.

Our results lend further support to the notion that ERP results from RSVP can be replicated in natural reading (see also Dimigen et al., 2011; Kretzschmar et al., 2009). The regression-contingent analysis shows that recovery processes are tightly coupled with regressions and the P600. The slow negativities in trials without regression may indicate a different processing strategy or the attempt to assign a sensible interpretation without revisiting earlier materials which leads to more effortful processing. These results also illustrate how coregistration can be used to uncover patterns that would go unnoticed in traditional experimental paradigms.

**Figure:** Average ERPs at electrode Pz from natural reading sessions.
Counterfactual conditionals provide two pieces of information: their literal meaning expresses a suppositional dependency between the antecedent and the consequent, their backgrounded meaning implicitly refers to the opposite factual state of affairs. Several event-related potential (ERP) studies have targeted the processing of counterfactual consequents, yet counterfactual antecedents have remained unstudied so far.

We present an ERP study which employed German past tense conditionals to compare subjunctive mood (which marks counterfactuality) with indicative mood (which marks hypothetical statements) at the critical point of mood disambiguation via auxiliary introduction in conditionals’ antecedents. In German this contrast is maximally parallel, as subjunctive mood introduction in past tense is accomplished by a minor variation of the auxiliary (e.g. waren vs. wären), while the rest of the sentence remains identical in counterfactual and hypothetical conditionals. Another advantage of German syntax is that the mood-disambiguating auxiliary takes the antecedent-final position where propositional content is already conveyed:

Counterfactual: Wenn die Bananen reif **wären**, dann wären sie grün/braun.

Hypothetical: Wenn die Bananen reif **waren**, dann waren sie grün/braun.

We collected EEG data of 16 subjects which read 52 counterfactual and 52 hypothetical conditionals that varied in semantic acceptability (2 × 2 design). At the critical first occurrence of the mood-disambiguating auxiliary we observed an early (200-400 ms) positivity followed by a transient (450-600 ms) negative deflection in frontal regions for subjunctive compared to indicative auxiliaries (see Fig.1). Importantly, no comparable differences between indicative and subjunctive were observed at the second occurrence of the respective auxiliaries, ruling out a merely perceptual effect of umlaut (a vs. ä). Further analyses show that linguistic mood did not interact with the N400/P600 truth-effect elicited by the unacceptable vs. acceptable sentence-final content words.

**Figure 1:** Grand average waveforms time-locked to the antecedent-final auxiliary disambiguating mood: Counterfactual (thick line) and hypothetical (thin line).

We discuss these novel findings in respect to working memory requirements for rule application and increased referential processing demands for the representation of counterfactuals’ dual meaning. Our result suggests that the counterfactual’s dual meaning is processed without any delay at the earliest point where counterfactuality is marked by subjunctive mood in the antecedent of a conditional.
ON THE IMPACT OF PHRASAL LENGTH ON NATIVE AND NON-NATIVE SENTENCE PRODUCTION: EVIDENCE FROM BASQUE AND SPANISH

Do native and non-native speakers resort to different mechanisms when choosing word orders? Hawkins (2014) argues that speakers tend to minimize the distance between the phrasal heads to render processing efficient (MiD principle) [1]. This principle predicts the emergence of different strategies depending on phrasal length and head-direction of the language: VO languages will favour short-before-long phrase orders and OV languages will prefer long-before-short phrase orders [2, 3, 4]. Previous findings in Korean (OV) indicate differences between native and non-native speakers [5]: English-dominant and balanced English-Korean bilinguals did not shift long elements before short ones, showing an overall preference for the short-before-long SOV basic word order. To further examine this issue, we ran two sentence production experiments to test whether Basque (OV) and Spanish (VO) non-native speakers differ from Basque and Spanish native speakers respectively in their sentence word order preferences.

In Experiment 1 (in Basque), we examined how highly proficient non-native speakers of Basque (n=24), who were native speakers of Spanish, differed from native speakers of Basque (n=24). In Experiment 2 we investigated how highly proficient non-native speakers of Spanish (n=18), who were native speakers of Basque, differed from native speakers of Spanish (n=18). Participants were presented a series of phrases of varying length (All-Short, Long-Subject/Long-Indirect Object and Long-Object) and were asked to arrange them in sentences (24 transitives, 30 ditransitives, 56 fillers). The results revealed a long-before-short preference in Basque in both transitive and ditransitive sentences and a short-before-long preference in Spanish only in ditransitive sentences, for both native and non-native speakers. Overall, non-native speakers showed a greater preference to use the canonical order than native speakers (significant in all conditions in Basque ditransitives and transitives and in only the long-O condition of Spanish ditransitives). Importantly, non-natives only showed a statistically smaller length effect in Spanish ditransitives (maybe due to their greater preference for canonical order), while similar length effects were shown by both groups in Basque transitive and ditransitive sentences.

In sum, for both native and non-native speakers, we found the opposite word order preferences for OV (long-before-short) and VO languages (short-before-long), as predicted by the MiD efficiency principle [1]. However, there was a difference with respect to the sensitivity to MiD in each group: non-native speakers showed a greater reliance on canonical word order. This might be related to the fact that it is the simplest sentence structure generated by the grammar [6]. The lack of shifting in English-Korean bilinguals can also be explained through this same preference for greater computation easiness [5]. Thus, our results indicate that a single processing mechanism (MiD) underlies the effect of phrasal length, which yields mirror word order patterns in VO and OV languages, regardless of being a native or non-native speaker, as well as a general non-native preference for syntactic simplicity over efficiency.

<table>
<thead>
<tr>
<th>Table 1. Percentages of shifted (non-canonical) sentences produced in Basque (transitive: O-S; ditransitive: O-IO) and Spanish (transitive: O-S; ditransitive: IO-O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitive sentences</td>
</tr>
<tr>
<td><strong>Basque</strong></td>
</tr>
<tr>
<td>Natives</td>
</tr>
<tr>
<td>3.6%</td>
</tr>
<tr>
<td>Non-natives</td>
</tr>
<tr>
<td><strong>Spanish</strong></td>
</tr>
<tr>
<td>Natives</td>
</tr>
<tr>
<td>Non-natives</td>
</tr>
</tbody>
</table>

We hypothesized that successful processing of sentences which mainly rely on the use of prosodic phrasing for correct interpretation is differentially affected by different noise maskers. The German structures in (1) and (2) can only be distinguished and correctly interpreted by the presence (1) or absence (2) of an intonational phrase boundary (IPB) after the first verb.

(1) [Ingo malt]_{IPB1} [Anna nicht]_{IPB2} aber beide essen ein Eis.
Ingo paints, Anna (does) not, but both are eating ice cream.

(2) [Ingo malt Anna nicht]_{IPB2}, aber beide essen ein Eis.
Ingo (does) not paint Anna, but both are eating ice cream.

Whereas speech recognition and processing are generally known to be negatively affected by background noise, acoustic cues used for prosodic phrasing (i.e., amplitude modulation, pitch contour, rhythm vs. F0 pitch contour) have been shown to be comparatively robust against a loud stationary noise masker (e.g., Carroll, 2013). Cooke’s (2006) glimpse model predicts better speech recognition in a noise masker with an amplitude modulated envelope compared to a stationary noise due to temporal release from masking. However, pilot findings suggested that the prosodic robustness observed for the processing of prosodic boundaries in stationary noise may not necessarily hold to the same degree in an amplitude-modulated noise masker, and may be affected differently than lexical recognition (Carroll, 2013). We propose that the naturally rhythmic amplitude modulation of the masker envelope may render prosodic cues (rhythm, duration vs. F0 pitch contour) less reliable during stream segregation. We tested the Closure Positive Shift (CPS; Steinhauer et al., 1999) as an electrophysiological ERP correlate of processing IPBs to determine the impact of loud stationary and amplitude modulated noise maskers on processing the prosodic information. When controlling for (near) perfect intelligibility, modulated noise could be presented at 4 dB louder than stationary noise. Twenty-five right-handed native German listeners between 20 and 30 years (\(\bar{x}\) 24.6) with normal hearing listened to 2 sentence structures as in (1) and (2) in three acoustic conditions, with 40 trials per condition. Electrophysiological brain responses (ERP) were measured at 64 scalp electrodes during listening. IPBs in stationary noise elicited a typical CPS effect comparable to a CPS elicited in quiet (360-800 ms), but delayed by about 100 ms (450-900 ms). Prosodic boundaries in the modulated noise also elicited a delayed CPS (450-900 ms), but in addition the CPS was reduced in amplitude. We interpret the time delay as an effect of increased listening effort due to general stream segregation mechanisms associated with speech in noise. The reduced amplitude of the CPS component suggests an additional processing difficulty associated with the segregation of the rhythmical amplitude modulated masker and the prosodically cued grammatical differences. Background noise seems to affect the processing of lexical items (or segmental information) differently than the processing of prosodic information, depending on the type of masking noise. Whereas rhythmic amplitude modulation allows for better speech intelligibility—in terms of louder masking noise at the same lexical intelligibility compared to stationary noise—it is more disruptive for the processing of suprasegmental information. Our findings bear implications for the interaction of bottom-up and top-down mechanisms, and underline the importance of testing language processing mechanisms in naturalistic conditions.

THE ROLE OF ORTHOGRAPHY IN RECOGNIZING CONNECTED SPEECH
Holger Mitterer (Università ta’ Malta) & Eva Reinisch (Ludwig Maximilian University Munich)
holger.mitterer@um.edu.mt

Several experiments suggest that orthography plays a role in spoken-word recognition (e.g., Rastle, McCormick, Bayliss, & Davis, 2011, JEP:HPP, 37, 1588-1594). The evidence for this assumption, however, is based on paradigms using single-word presentation. The present study hence addressed the role of orthography in listening to connected speech, since, here, the ubiquitous reductions (e.g., yesterday → [yeshay]) lead to massive inconsistencies between spoken form and orthographic representation. This makes an involvement of orthography when listening to natural connected speech a-priori less likely.

This possibility was investigated by looking at the processing costs of deleting speech sounds that are orthographically coded or not. If orthography is involved in recognizing connected speech, reduction of orthographically coded sounds should be more detrimental than reduction of sounds that are not coded in the orthography. The example used here is the glottal stop in German, which occurs in the canonical pronunciation of orthographically vowel-initial words (e.g., Affe, Engl., ape, = /ʔafaf/). We compared the reduction costs of deleting this glottal stop with the deletion of two orthographically coded segments; in Experiment 1, with the deletion of the glottal stop in a language where it is orthographically coded (Maltese, where q is the letter for /ʔ/, e.g., qamar, /ʔamar/, Engl., moon), and in Experiment 2 with the deletion of another German phoneme, /h/ (also written as “h”).

Experiment 1 made use of the word-monitoring task: Participants see a written target word on the screen and have to press a button if this word occurs in the speech stream. /ʔ/-initial words were presented after words or nonwords (e.g., target ape: sunape versus munape) and with the glottal stop present of not. The lexical manipulation evaluated the status of the German glottal stop, which some assume to be a phonetic boundary marker. Phonetic cues to word boundaries have little leverage if there are lexical cues (such as a preceding word, see Mattys & Melhorn, 2007, JASA, 122, 554-576). If the German glottal stop is just a phonetic boundary marker, it should only have an effect when the preceding context is a non-word, while the presence versus absence of the glottal stop in Maltese should have a less restricted effect. However, the results were similar across languages: Stimuli without glottal stops led to slower responses but only a few misses in both languages. This effect was as strong as the lexical effect and independent of it (i.e., two main effects, no interaction).

Experiment 2, using visual world eye-tracking, compared the reduction costs for /ʔ/ and /h/ in German. Participants heard slightly constraining sentences such as Beim Skifahren hat er immer einen Helm auf (Engl., When skiing, he always wears a helmet) and saw visual displays containing a target (helmet), another semantically fitting item (backpack), and a distractor (e.g., a baboon). Participants were asked to click on a picture if mentioned in the sentence. By cross-splicing, two versions of the sentence were generated: with and without the initial /h/ of Helm (and for /ʔ/-initial items, with the /ʔ/ present or not). To allow participants to reject reduced items, 20% of the trials were fillers with unrelated displays and sentences. Results showed similar reduction costs for the orthographically coded initial /h/ and the orthographically absent /ʔ/ in accuracy (lower if the segment is absent), reaction times (longer if the segment is absent), and eye-movements (slower to converge on the target).

These results lead to two conclusions. First, the role of orthography in spoken-word recognition seems to be reduced in connected speech. This is most likely due to the mismatch between an invariable orthographic code and the variable acoustic forms of spoken words. Second, the results show how strongly our thinking about—but not our processing of—speech is influenced by learning to read. Most speakers of German are completely unaware of the glottal stop in German. Yet they produce and perceive this segment just as any other segment that is coded orthographically. This is in line with evidence that processing of speech in terms of segments is not tied to learning to read (McQueen, Tyler, & Cutler, 2012, LLD, 8, 317-339).
WHAT WAS THE QUESTION? BROAD FOCUS IN FOCUS-SENSITIVE COORDINATION

Jesse A. Harris (UCLA) & Katy Carlson (Morehead State University)
jesseharris@humnet.ucla.edu

Processing some temporarily ambiguous ellipsis structures, like gapping, seems to be guided by preferences favoring less complex structure (Carlson, 2002). However, focus-sensitive coordination (FSC) structures with let alone or much less are argued to require ellipsis (Hulsey, 2008; Toosarvandani, 2010), and we propose that mandatory ellipsis removes any pressure for a syntactically simpler alternative, on the assumption that the cost for retrieving the correlate for ellipsis is constant no matter the size of the remnant (Frazier & Clifton, 2005; Martin & McElree, 2008). Indeed, Harris (2013) found that completion of FSC fragments like (1) showed a bias towards the more complex VP (1a) over the simpler NP (1b) alternative. In two experiments and a corpus study, we provide evidence that the processor prefers remnants that limit accommodating unsupported discourse commitments (see also Kim et al., 2012). Specifically, we suggest the constraint in (2) favoring the widest possible Question Under Discussion – i.e., the question that corresponds to the most immediate topic of the discourse (QUD; Roberts, 1996). This predicts that structurally more complex remnants corresponding to broad focus, such as VPs, may be preferred in isolated sentences, which may require the comprehender to infer an appropriate context.

In Experiment 1, a written forced-choice completion experiment, 30 subjects selected one of the following options to complete items like (3): a. VP with pronoun (47%), b. VP with new noun (17%), or c. novel NP (36%). These varied both the syntactic structure and the location/size of contrastive material. The 64% total bias toward either VP structure replicates Harris (2013), while the preference for (a) over (c) shows a preference for only one contrastive constituent following let alone. Choosing a VP over an NP keeps the QUD as broad as possible to avoid unnecessary discourse commitments (2).

In a study of the British National Corpus (BNC), we found a slight preference for NP (47%) over VP (37%) remnants. A similar distribution was observed in the Contemporary Corpus of American English for let alone as well as much less constructions. These distributions contrast with the VP bias in questionnaire studies (Harris, 2013), suggesting that (i) the VP bias for FSCs in Experiment 1 is not syntactic in origin, and (ii) the bias cannot be attributed to the frequency of the structures. Within the corpus, these sentences are found in contexts where little to no accommodation of the QUD is needed; thus, the Minimize QUD principle in (2) does not apply, as it predicts a VP bias through broad focus only in the absence of contextual bias.

If likely discourse structure guides expectations about the contrast in FSCs rather than syntax (see Clifton & Frazier, 2012), then available discourse-related information such as contrastive pitch accents should facilitate processing of remnants congruent with the implied QUD. In Experiment 2 (auditory forced-choice completion), 48 subjects chose between VP and NP remnants after either NP (4a) or V (4b) focus in the correlate. As predicted, pitch accent location strongly affected remnant choice. Following NP focus (a), NP remnants were preferred (61%); following V focus, VP remnants were preferred and NPs dispreferred (28% NPs). Thus FSCs are clearly sensitive to changes in information structure and the implicit QUD as signaled by focus (Toosarvandani, 2010).

In sum, language users use default interpretations of the discourse to guide their expectations about what FSCs will contrast with, minimizing potentially unnecessary discourse commitments, rather than using purely syntactic or exposure-based preferences.

(1) The patient didn’t even eat dinner let alone … {a. make it | b. dessert}
(2) Minimize QUD: Default to the broadest QUD possible, unless there is evidence from prior discourse to the contrary.
(3) Simon couldn’t remember the attack, let alone …
   {a. describe it | b. describe the mugger | c. the mugger}
(4) a. John doesn’t like MARY, let alone … | b. John doesn’t LIKE Mary, let alone …
   {Sue | love her}
The relevance of intonational phrase (IPh) boundaries to sentence processing is well-established, but recent studies have yielded mixed results for intermediate phrase (ip) boundaries. It has been suggested (Millotte et al.) that this is because ip boundaries are not naturally employed by speakers if they divide a sentence into inappropriately short phrases. This is in line with the theoretical linguistics literature on the syntax-prosody interface (Selkirk 2000) which notes an interplay between optimal phrase lengths and structural constraints on prosodic phrasing. Our performance data confirm the relevance of phrase lengths in the processing of a double-PP construction in Hebrew.

Materials and method: Selection restrictions force PP1 to attach low, but PP2 attachment is ambiguous: it can attach maximally high to VP (as an argument of hid) or maximally low to the NP inside PP1 (modifying divorce). PP2 was either short (as shown) or long (with an added modifier). The subject pronoun had its antecedent in a semantically neutral context sentence.

\[ \text{hu } \text{VP} \text{[histir } \text{NP} \text{[et ha-ka\'as šelo } \text{PP1[ al ha-geru\'šin } \text{PP2[ mi-tami]]}} \]

This double PP construction affords more potential locations for prosodic boundaries, and has a sharper structural contrast between the two attachment sites, than the much-studied single-PP construction for which conflicting results have been found.

Participants read a context+target sentence pair aloud and then chose between two interpretations (divorce from Tami / hid from Tami).

Statistical analyses of ear judgments and attachment data were conducted with Multilevel Cross-Classified Multinomial Logistic Regression Models.

Results (prosody): Ear judgments and acoustic analyses showed significantly more instances of a pre-PP2 ip boundary for long PP2 (57.8%) than for short PP2 (17.8%). As anticipated, short PP2 items favored breaks before PP1 (34.0%) or no breaks at all (41.5%).

Results (ambiguity resolution): There was a significant effect of phrase length on preferred interpretation (55.9% high attachment for long-PP2 targets; 46.2% for short); also a significant effect of pre-PP2 breaks on interpretation regardless of phrase lengths: high PP2-attachment was favored by an ip boundary immediately preceding PP2 (59.5% high for long-PP2; 60.7% for short), with significantly less high attachment for pre-PP1 breaks and No Breaks. A multi-level structural equation model analysis of the direct and indirect effects between PP2 length, break location and attachment height found a direct effect of length on attachment but also a significantly stronger effect of break location on attachment (Wald’s Chi-Squared, one-tailed=3.26, p=.035).

Interpretation: This study thus provides support for the assumption that phrase length effects on attachment preference are mediated at least in part by the prosody that is projected by the reader, as has often been hypothesized in the ‘implicit prosody’ literature (Fodor 2002) but which can be difficult to establish.

References:

Attachment is a fundamental part of syntactic processing, as it involves figuring out how to link upcoming words with the current syntactic structure. We know that a prosodic boundary separating a phrase from the nearest attachment site favors high attachment (Watson & Gibson 2005, Carlson et al. 2001). Also, relative clause attachment can be affected by accenting the potentially modified noun (Schafer et al. 1996, Lee & Watson 2011). But it’s not clear whether accents affect attachment ambiguities more generally, and whether their effects depend on prosodic boundaries. Three auditory questionnaires accented the higher or lower attachment site in different structures, and found that accents drew the attachment of an ambiguous phrase without depending on boundaries. These results show that accents affect more than just relative clause processing, demonstrating a need to incorporate accents into general theories of the role of prosody in sentence processing.

Experiment 1, an auditory questionnaire (N=72), contained 20 items like (1), with a final phrase that could modify the higher or lower verb. Sentences were recorded in four conditions: with or without a large prosodic boundary (IPh) before the final phrase, and with a L+H* pitch accent on Verb1 (claimed) or Verb2 (lied). Listeners chose between two visual paraphrases to indicate their interpretation of the sentences. Accenting Verb1 led to a significant rise in high attachments vs. accenting Verb2 (p’s<.05); the presence of the prosodic boundary also led to more high attachments (p’s<.05), but no interaction.

Experiment 2, an auditory questionnaire (N=48), contained 24 sentences like (2), with a final phrase that could modify the higher verb or the verb inside a relative clause. Sentences were recorded in four conditions: with a smaller or larger prosodic boundary (ip or IPh) before the final phrase, and with a contrastive accent on Verb1 (comforted) or Verb2 (insulted). Accenting Verb1 produced more high attachments vs. accenting Verb2 (p’s<.05); prosodic boundary size had no significant effect and no interaction with accent. Carlson et al. (2009) showed a similar effect with modified pitch contours, but this shows it in naturally recorded speech.

Experiment 3, an auditory questionnaire (N=48), tested 20 sentences like (3), with a final PP that could attach to the verb or the object noun. Sentences were recorded in four conditions: with or without a small prosodic boundary (ip) before the final phrase, and with a contrastive pitch accent on the Verb (entertained) or the object Noun (toddler). Accenting the Verb led to more verb attachments vs. accenting the Noun (p’s<.05); the presence of the prosodic boundary also increased high attachments (p’s<.05), with no significant interaction.

The finding that accents affect attachment in several different structures (PP or RC attachment, to verbs or nouns) challenges the usual view of the role of pitch accents in processing. They usually indicate emphasis and the position of focus within a sentence (Rooth 1992), allowing a sentence to fit into a larger discourse context. These results suggest that accents, and perhaps focus, can also influence basic syntactic structure. Further, the effects of accents do not seem to depend on the prosodic phrasing of a sentence, leading to interesting questions about the overlapping roles of the two types of prosodic units.

1. Paula claimed that Alex had lied # on Monday.
2. Jimmy comforted the girl that he had insulted # after the party.
3. Alison entertained a toddler # with many toys.
In English (amongst other languages) the RC attachment ambiguity has been employed to study the constraints that underlie language processing. In sentences like Someone shot the servant of the actress who was serving tea, plausibility constrains the attachment of the relative clause to either servant (high attachment) or to actress (low attachment) (e.g., van Gompel et al., 2005). When plausibility cues are absent as in Someone shot the servant of the actress who was old, it has been shown (Clifton et al. 2002) that the presence (HA cue) or absence (LA cue) of a pause before the relative clause aids participants’ resolution of the impending syntactic ambiguity. However, it has also been shown that there is a preferred resolution of the ambiguity, i.e. low attachment and high attachment is the dispreferred attachment option.

In 2 experiments we investigated how prosody and plausibility interact in resolving this syntactic ambiguity taken into account a possible LA preference. Using a syntactic priming paradigm (cf. Scheepers, 2003) the first of two experiments (72 participants) looked at how plausibility and prosody (in a spoken prime sentence like 1a-d) concurrently influence RC-attachment preferences in a subsequent target sentence fragment for completion (e.g. The tourist guide mentioned the bells of the church that …).

(1) Someone shot the servant of the actress
   a. who was serving tea.
   b. who was very famous.
   c. [Pause] who was serving tea.
   d. [Pause] who was very famous.

Binary logistic analyses using Generalised Estimating Equations (GEE) revealed no main effects of or interactions with either predictor (prosody or plausibility) across participants or items. However, after a subsequent cluster analysis (using the squared Euclidian distance and Ward’s method) returned two groups of participants (32 v. 40 participants) with opposing effect patterns, GEE was run again for each cluster. Analyses for cluster 1 (32 participants) showed a prosody*plausibility interaction (all p’s < .001). The two no-pause conditions (1a and b) did not show any appreciable priming effects. In contrast, in the two pause conditions showed reliably more HA target completions when the prime-RC was semantically biased towards high - (1c) than towards low-attachment (1d). Cluster 2 (40 participants) also showed a clear, but very different prosody*plausibility interaction (all p’s< .001): Again, no priming effects were found for the two no-pause conditions (1a and 1b), but the comparison between the two pause conditions (1c and 1d) now showed reliably fewer HA target completions when the prime-RC was semantically biased towards high (1c) than towards low-attachment (1d). In sum, two clusters were identified that displayed opposing priming results especially in the pause conditions. The first cluster showed the strongest priming effect when both dispreferred cues to high attachment (pause + semantic support for HA) agreed in their attachment. The second cluster showed the strongest priming effect when the cues disagreed in their attachment support (pause + semantic support for LA).

The second experiment was an exact replication of the first, but using a new set of participants (72). Again, a cluster analysis identified two clusters (39 v. 33 participants) that displayed opposing priming effects especially in the pause conditions. Moreover, the second cluster showed strong priming effects in the no pause conditions, i.e. more HA completions after semantically biased HA than after semantically biased LA primes.

The experiments showed that the two types of cues interact in a complex way. Three processing principles that could help explain the different priming results are proposed. That is, the interaction of plausibility and prosody could be governed by:

1. The prominence of the prosodic high attachment cue (pause), 2. The surprisal associated with the dispreferred high attachment cues (pause v. semantic HA), 3. The type of structural revision necessary.
A SKETCH OF AN EXTRINSIC TIMING MODEL OF SPEECH PRODUCTION

Alice Turk (University of Edinburgh), & Stefanie Shattuck-Hufnagel (Massachusetts Institute of Technology).
turk@ling.ed.ac.uk

Articulatory Phonology/Task Dynamics (AP/TD, Browman & Goldstein 1985, Saltzman & Munhall 1989 et seq.) is the model of speech production that currently provides the most comprehensive account of speech timing phenomena. Timing control in this model is intrinsic, that is, surface timing patterns emerge from properties of the system and do not need to be specified, tracked or modified during an utterance. However, several lines of behavioral evidence challenge intrinsic timing as implemented in Articulatory Phonology/Task Dynamics, and support the view that timing control in speech production is extrinsic. These include 1) separate control of movement onsets and offsets, difficult to implement in mass-spring systems such as AP/TD 2) increasing variability with increases in interval duration, as predicted by a "noisy timekeeper" model, 3) language-specific constraints on surface timing in a quantity language, not predicted in intrinsic timing systems which do not allow surface timing specifications, and 4) motor equivalence of strategies for producing surface duration patterns, again difficult to explain in intrinsic timing systems with no reference to surface time. These lines of evidence motivate the consideration of alternative models. In this paper, we present a sketch of an alternative extrinsic timing model of speech motor control. This approach involves three distinct phases: 1) Phonological planning, 2) Phonetic planning, and 3) Motor-sensory implementation. The phonological planning stage involves structuring symbolic segmental representations into a hierarchy of prosodic constituents and prominences. We assume that phonetic planning involves balancing task requirements and movement costs to yield (near-) optimal parameter values (cf. Optimal Control Theory approaches) for use in the third, Motor-sensory implementation stage. Task requirements include things like being accurate, spreading information evenly throughout the signal via an appropriate prosodic structure, and not taking too long. Movement costs include things like the spatial inaccuracy cost of moving fast, and energy expenditure. The phonetic planning stage involves planning a sequence of goal states (e.g. spectral properties that can serve as cues to planned distinctive features), the timing between goal states, the articulators that create the goal states, spatial goals of their movements, and movement timing characteristics, including the timing of movement onsets and movement time course characteristics. In this approach, articulatory overlap results from movement goals which follow each other in (relatively) rapid succession. As speech unfolds in the Phonetic implementation stage, we assume that speakers continuously track their movements to reach their targets with desired spatial and temporal accuracy.

In addition to presenting our model sketch, we discuss available model components that could be used to implement it. These include Optimal (Feedback) Control Theory (e.g. Todorov & Jordan 2002), DIVA and VITE (Guenther 1995, Bullock & Grossberg 1988), General Tau theory (Lee 1998), and timekeeper models (e.g. Gibbon 1997).

References


Sentence formulation involves encoding of syntactic structure and phonological form. Models of language production assume that a conceptual structure is first mapped onto a syntactic representation (grammatical encoding). The prosodic-phonological representation is specified on the basis of the syntactic representation. Only few studies suggest a prosodic influence on grammatical encoding during production [e.g., 1, 2].

Two studies reported here show that speakers and readers consider prosodic phrasing when choosing among syntactic options. Specifically, the use and interpretation of causal adjuncts in the context of a negation (1) were investigated. In German, causal adjuncts may be expressed as sentential adverbs (weil-CP) or as prepositional phrases (wegen-PP). While sentential adjuncts typically project their own intonation phrase (IP), PPs are prosodically integrated. This prosodic difference may be crucial when the causal adjunct is negated. In the case of a negated weil-CP (1a), a prosodic break preceding the weil-clause (i.e., the left edge of the corresponding IP) may separate the negation from its scope [3], and lead to erroneous association of the negation to the matrix clause. If speakers factor in prosodic phrasing when choosing a syntactic construction, the negated adjunct should preferably be expressed by means of a prosodically integrated PP (1b).

To test this hypothesis, we searched the Parlament-corpus (Cosmas II) for tokens of the word sequences [nicht, weil] (not because-CP) and [nicht wegen] (not because of-PP) and, for the first 500 tokens of each, determined the scope of the negation. In 41% of the [nicht, weil]-CPs, the negation scopes over the causal adjunct. The same scope condition is found in 91% of the [nicht wegen]-PPs. This highly significant difference attests the speakers’ consideration of prosodic phrasing during sentence construction and, correspondingly, suggests strong interaction of syntactic and prosodic encoding.

A subsequent rating study (45 participants) with 24 written stimuli like (1) yields a significant interaction of scope of negation and syntactic construction (PP vs CP) in reading. Sentences with a strong pragmatic bias for the negation scoping over the causal adjunct (1ab) are generally rated worse than non-negated adjuncts (1cd) (replicating findings by [4]). More importantly, the devaluation is significantly stronger for negated CPs (1a) as compared to negated PPs (1b). Arguably, the comma, and the corresponding implicit prosodic boundary [5] that readers project between the negation and the weil-CP, hampers computation of the required scope. In the case of negated PP-adjuncts, there is no intervening comma (and no corresponding prosodic break) that would prevent proper computation of the scope.

(1) Peter kauft das Auto nicht...
Peter doesn't buy the car...
   a. ..., weil es günstig ist.        c. ..., weil es teuer ist.
   because it is low-priced.        because it is expensive
   b. ... wegen des günstigen Preises.
   because of the low price.
   b. ... wegen des teuren Preises.
   because of the high price.

Nb. (1ab) implicate that P. does buy the car for another reason than the low price

To the extent that both speaking and reading involve sentence formulation (the latter on the basis of the written string, the former on the basis of conceptual structure), these results strengthen the case for an interaction of grammatical and phonological encoding.

References:
The ability to develop spelling-to-sound correspondences depends on the regularity of the language spelling sound correspondences, which may affect the grain size of the reading units and the role of phonological information (e.g., Ziegler et al., 2010). Italian is very regular, except for the assignment of stress to multisyllabic words, as stress is unpredictable and conventionally assigned. Moreover, most Italian multisyllabic words are stressed on the penultimate syllable (e.g., geLAto, ice cream; i.e., dominant stress), whereas only a smaller proportion (~ 20%) bears antepenultimate stress (e.g., TAvolo, table). Therefore, while stress can be correctly assigned only through lexical look-up, stress dominance (the pattern of the most frequent stress) would encourage readers to assign penultimate stress, as a sort of regularity effect (e.g., Colombo, 1992).

In the present work we investigated the acquisition of stress assignment in Italian. To this aim, we examined the frequency effect, as a marker of the use of lexical knowledge in assigning stress, together with stress dominance, as a marker of overall distributional information of stress patterns in Italian, implicitly learned from spoken language. We tested second- and fourth-graders in a reading-aloud experiment, in which we included high- and low-frequency words with a strong dominant or non-dominant stress neighbourhood, and nonwords, the latter constructed with a high probability to be pronounced with the dominant (e.g., miffone) or the non dominant (medulo) stress pattern.

The analyses on words stimuli showed three interesting results: a) word frequency affected the reading of both second- and fourth-graders; b) children of both age groups made more stress errors when reading antepenultimate than penultimate stress words, but only when words were of low-frequency; c) the latter trend was more pronounced in second- than fourth-graders. The analyses on nonwords stimuli showed that second-graders tended to assign stress mostly according to stress dominance (i.e., most nonwords received dominant stress) and the probability of a nonword receiving stress according to its stress neighbourhood (information associated with the ending of words and nonwords) increased with age.

Our results shed light on important issues regarding the acquisition of reading in a very regular language like Italian. First, the presence of the frequency effect in beginning readers (second-graders) suggests that lexical knowledge is important even in a transparent language and is exploited from the early stages of reading development. The presence of a larger number of stress errors in second- than in fourth-graders suggests however that their ability to use lexical information was limited, relative to the older children. Second, with regard to stress assignment, the results obtained with words and nonwords help to outline the developmental trajectory of the assignment of stress: Besides lexical information, children learn to use two types of distributional information to assign stress to unknown stimuli in reading - that is, knowledge of the overall relative distribution of stress, and specific information about direct links between orthography and stress from word endings (stress neighbourhood). These two types of knowledge seem to have different developmental trajectories, with general information about stress being acquired earlier than stress neighbourhood. Our findings also show that both item-specific knowledge and general information about stress distribution are relevant in children's reading, suggesting the simultaneous use of both lexical and sublexical information.

References
WHEN EXACTLY CORNERS STOP CORNING?
INCREMENTAL MASKED PRIMING AND COMPLEX WORD IDENTIFICATION
Davide Crepaldi (University of Milano Bicocca), Marco Marelli (University of Trento), Simona Amenta, Anais De Santis, Elena Angela Morone & Rocco De Marco (University of Milano Bicocca)
davide.crepaldi1@unimib.it

It is widely accepted that complex words such as “dealer” are decomposed into their constituent morphemes during visual word identification. This process seems to be unaffected by whether morphemes actually contribute to the word meaning – it would also break down “corner” into “corn” and “er”. However, results from classic masked priming studies are not completely clear-cut (e.g., Davis and Rastle, 2011). Here we address this issue using incremental masked priming, which: (i) allows to track the temporal pattern of the effects by varying quasi-continuously the Stimulus Onset Asynchrony (SOA) between primes and targets within the same experiment; (ii) gets rid of unrelated primes by using the shortest-SOA condition as a baseline; (iii) retains the critical feature of classic masked priming, i.e., primes are kept outside participants’ awareness, by using only sub-threshold SOAs (<60 ms).

In Experiment 1, 95 prime-target pairs of Italian words were devised in each of 3 conditions. In the first, genuine derived words primed their stems (artista-ARTE, artist-ART); in the second, pseudo-derived words primed their pseudo-stems (retaggio-RETE, legacy-NET); in the third, simple words primed orthographically related stems (corallo-CORO, coral-CHOIR). Prime-target pairs were tested in five different SOA conditions: 12ms, 24ms, 35ms, 47ms, and 59ms. Data were analyzed using mixed-effects models; only statistically supported results are reported below.

Data on 208 participants show that: (i) genuine derivations and pseudo-derivations have similarly-shaped priming trajectories across SOAs; (ii) when compared with each other, they are statistically equivalent in size up to the 47-ms SOA; (iii) when compared individually to the orthographic baseline, pseudo-derivational priming emerges earlier than genuine derivational priming (35-ms SOA vs. 47-ms SOA).

Experiment 2 is a replication of Experiment 1 in a new sample of 141 participants. By using higher-frequency monitors, we were able to refine our “temporal resolution” and test seven different SOAs (8, 17, 25, 33, 42, 50, and 58 ms). Results were closely similar to those of Experiment 1 and revealed again: (i) similar priming trajectories across SOAs for dealer-DEAL and corner-CORN; (ii) equivalent priming for these pairs at all SOAs; and (iii) earlier significant priming for pseudo-derivations (SOA=25 ms) than for genuine derivations (SOA=33 ms).

Experiment 3 is a validation of the paradigm that was carried out using unrelated primes (and the same seven SOAs tested in Experiment 2). As expected, no consistent priming pattern was observed, thus confirming that data in Experiment 1 and 2 were driven by the nature of the relationship between primes and targets.

These data confirm that morphological facilitation is equivalent for genuine derivations and pseudo-derived words at the earliest steps in visual word identification. However, they also point to a more intricate interaction than what previously thought between form and semantics, as revealed by the surprising earlier effect with pseudo-derivations than with genuine derivations.

Reference
DISTRIBUTIVE AND COLLECTIVE READINGS OF PLURAL DPS
Carmen Dobrovie-Sorin (CNRS, Paris Diderot), Emilia Ellsiepen (Universität Frankfurt) & Barbara Hemforth (CNRS, Paris Diderot)

We will report the results of four experiments that were meant to investigate the availability of distributive readings in three types of sentences, which all involved singular indefinite DPs in the object position but differed wrt to the syntax of the subject DP (DP_{def-pl}, groupDP, the majority of DP):

1. **Les enfants** ont construit un château de sable. Finalement, (a) le château/ (b) les châteaux étaient fragile(s). The children built a sand castle. Finally, (a) the castle/ (b) castles was/were fragile.

2. **Le groupe d’enfants** a construit un château de sable. Finalement, (a) le château/ (b) les châteaux étaient fragile(s). The group of children built a sand castle. Finally, (a) the castle/ (b) castles was/were fragile.

3. **La plupart des enfants** a construit un château de sable. Finalement, (a) le château/ (b) les châteaux étaient fragile(s). The majority of the children built a sand castle. Finally, the castle/castles was/were fragile.

We ran three acceptability judgment experiments with French examples as in (1a-b, 2a-b), varying DP type (1: DP_{def-pl}, 2: groupDP) in the first sentence and a singular (a) vs. plural (b) anaphoric definite DP in the second sentence (respectively requiring a collective vs. distributive reading). Participants judged the experimental items on a scale from 1-10. All experiments were run on Ibex Farm (Alex Drummond; http://spellout.net/latest_ibex_manual.pdf) with voluntary native French speakers. In Experiment 1, 40 participants were presented with 24 experimental items (6 per condition) following a Latin Square design mixed with 32 fillers. Our results (analyzed with maximal structure mixed linear models) showed a strong preference for the collective reading (i.e., singular referents in the second sentence were preferred, p<.01), but no interaction between DP-Type (DP_{def-pl} vs groupDP) and collective vs. distributive reading (ps >.5). To increase chances for an effect of DP-Type, we exchanged some of the experimental items, which seemed to be somewhat biased for collective/distributive readings across DP-Types for Exp. 2 (77 participants). We obtained the same pattern of results as Exp. 1, except that in the first few items presented to the participants there seemed to be a numerical (non-significant) hint of an interaction of DP-Type and number marking on the DP in the second sentence. To exclude that the lack of an effect of DP-Type might be due to experimental artifacts (priming, filler/target sentence ratio), we reduced the number of experimental items for Experiment 3 to 16 critical items and increased the number of fillers to 48. Collective readings were still judged significantly more acceptable than distributive readings (ps < .01) and we did not find any evidence for an interaction between DP-Type and collective/distributive readings (p>.5; DP_{def-pl} vs groupDP; p=.01). Experiment 4 investigated whether distributivity is made possible/facilitated by plural quantification. In this experiment, we varied DP type (DP_{def-pl}, 1, vs. the majority of DP, 3) and singular vs. plural definite DP (le (a)/les(b) château/x) in the second sentence as before. Acceptability judgments (33 participants, 16 items) showed a reliable interaction (p<.01) between DP-Type and distributive vs. collective readings with collective readings being judged more acceptable for plural DPs (coll: 9.34; distr: 6.75) as before but distributive readings being judged more acceptable for the majority of DPs (coll: 6.9; distr: 8.1).

In line with Dotlacil et al., (2010), our results contradict the view that the distributivity of plural definite DPs is generated by the pluralization of the VP triggered by the plural marking on DP arguments (Kratzer 08). Distributivity is dispreferred with plural definite arguments as confirmed by but the absence of a contrast with groupDPs in 3 experiments, but can be triggered by the majority of DPs. We will argue that our data point to the necessity of a quite drastic revision of prevailing theories of plural predication and plural quantification.

BETWEEN-GROUP MATCHING OF CONFOUNDING VARIABLES: WHY COVARIATES REMAIN IMPORTANT FOR ANALYSIS

Christoph Scheepers (University of Glasgow)
christoph.scheepers@glasgow.ac.uk

Experimental designs are often within-subjects and between-items, or vice versa. For example, in lexical decision experiments, conditions frequently involve different types of words (between-item manipulation) all of which are presented to each participant (within-subject manipulation). Such designs almost always pose the problem of controlling for confounds (lexical frequency etc.). Ideally, the different groups of items should not differ on such control variables. However, since a ‘perfect match’ is often hard to achieve, researchers typically rely on performing a significance test at item level (e.g., do Type-A words differ from Type-B words in terms of lexical frequency?) assuming that in case of a non-significant difference (say, $p > .1$) they can “rest assured” that the confounding variable is unlikely to explain any effects of experimental condition (e.g., type of word). I will show that one can actually not “rest assured” in this case, and that the currently perhaps most common practise of dealing with between-group confounds in such designs is flawed: Performing a significance test at item-level does not take into account that the same set of items is repeatedly presented to many participants. The latter increases the effective sample size (and therefore power) for any cross-condition bias in the confounding variable, and thus, any influences such a bias might have on the experimental results.

To illustrate this, I will present results from a series of Monte Carlo simulations. Each simulated within-subject/between-item experiment had two conditions (20 items each) and up to 30 “subjects”. The two groups of items always differed non-significantly ($p$ ranging from .10 to .97; $t$-test) on a randomly determined covariate with known linear effect on the dependent variable of interest. The generative model added various sources of random variation at subject, item, and trial level to the data (cf. Barr, Levy, Scheepers, & Tily, 2013), and assumed either no effect of experimental condition (estimating Type I error rate) or a condition-effect in the same or the opposite direction to the bias in the covariate (estimating power). Data analyses were based on $F_1$ AN(C)OVA and maximal Linear Mixed Effects Models (LMEMs), each time either including or excluding the confounding variable as a covariate. The results were clear. Ignoring the covariate led to unacceptably high Type I error rates in $F_1$, with more subjects leading to increasing levels of anticonservativity; conversely, power was greatly reduced for condition effects that were masked by an opposing bias in the covariate. Including the covariate in the analysis alleviated both of these problems, but not completely because $F_1$ does not account for additional by-item random intercept variance in the generative model. Maximal LMEMs without covariate were too conservative when condition effects were masked by an opposing bias in the covariate (their performance generally suffered from conflating covariate-related item variance with additional by-item random intercept variance in the generative model). Maximal LMEMs including the covariate were near-optimal in terms of both Type I error rate and power.

To conclude, showing a non-significant bias at item-level is actually not very useful when dealing with confounding variables like (say) lexical frequency, particularly when the same set of items is presented to many participants. Indeed, such control variables should always be included in the analysis model. Note that the labels ‘subject’ and ‘item’ are fully interchangeable—the exact same problem occurs when groups of subjects (assumed to be matched on some person-specific confound) are repeatedly tested over many items or trials.
A recent eye-tracking reading study showed that reading times for sentences expressing similarity between two ‘and’-coordinated abstract nouns (e.g., ‘Talent and wisdom are certainly similar...’) can be modulated by non-referential spatial information (the distance between two playing cards) preceding the sentence (Guerra & Knoeferle, 2012). These findings suggest that, just as for concrete nouns and verbs, non-linguistic visual information can influence interpretation of abstract adjectives. Furthermore, they imply that non-linguistic information can affect sentence processing in the absence of an overt referential or lexical-semantic links. Yet, a number of issues remain open regarding the mechanisms driving the interaction between non-referential spatial information and abstract language processing. Guerra and Knoeferle relied on a co-indexing mechanism, which relates corresponding elements in the sentence and in the visual context. Previous research (Knoeferle & Crocker, 2007) showed that, in referential visual context, concrete nouns and verbs co-index with visually depicted objects and events, respectively. However, in non-referential visual contexts it is less clear how abstract linguistic representations co-index with visually depicted spatial distance. In two eye-tracking reading studies, we assessed the linking hypothesis between non-referential visual information and abstract language. The goal of these studies was to clarify whether co-indexing occurs beyond a one-to-one mapping of nouns to objects.

A 2x2 Latin square design combined spatial distance (playing cards close vs. far apart) with the notion of social relations (friendly vs. unfriendly) expressed by abstract adverbs of manner. In both experiments, participants examined a visual context showing two playing cards (that moved either far apart or close together), followed by a German sentence describing a friendly (vs. unfriendly) encounter. Using different syntactic structures, we examined how and when abstract language co-indexes with spatial distance. In Experiment 1 (see 1., literal translation of the German original), we asked if spatial distance could modulate abstract interpretation even when there is no one-to-one mapping between coordinated objects and coordinated nouns. If so, then effects of distance should be time-locked to the sentence region expressing the manner of the social relation, or the immediately ensuing region (cf. Guerra & Knoeferle, 2012). In Experiment 2 (see 2.), we examined whether spatial distance could modulate sentence interpretation even before both sentential nouns can be co-indexed with the objects in the visual context. If so, spatial distance effects should appear as soon as the critical adverb expresses the manner of the social relation.

1. ‘Sandra’s met VP her aunt NP2 cheerfully/grumpily ADV in NP the elevator NP3’.
2. ‘Sandra’s met VP cheerfully/grumpily ADV her aunt NP2 in NP the elevator NP3’.

In Experiment 1 we found faster reading times when sentences expressing friendly relations were preceded by cards close together (vs. far apart), and the opposite pattern for sentences that expressed unfriendly relations. This interaction was reliable in first-pass and regression path duration at the region immediately after the critical adverb (PP). In Experiment 2, a reliable interaction effect was found in first-pass at the critical adverb, where longer reading times emerged for sentences expressing friendly relations when they were preceded by cards close together (vs. far apart), and vice versa for sentences expressing unfriendly social relations.

Our findings show that abstract language can be co-indexed with spatial distance as soon as critical information is available for the reader, and beyond concrete nouns and verbs. Additionally, they suggest that spatial information-abstract language mapping can speed-up (Experiment 1) or slow-down (Experiment 2) incremental semantic interpretation. We discuss the implications of these results for current account of situated language processing.
THE EFFECT OF SOCIAL CIRCLE SIZE AND HETEROGENEITY ON SEMANTIC SKILLS

Shiri Lev-Ari (Max Planck Institute for Psycholinguistics)
shiri.lev-ari@mpi.nl

People differ in the size and heterogeneity of their social circle. Previous research shows that people learn phonetic features better if they are exposed to more variable input (e.g., Bradlow & Bent, 2008). These experiments test whether having a larger or more variable social circle, and therefore more variable input, boosts semantic skills, and, specifically, lexical prediction (exp. 1) and comprehension of evaluative expressions (exp. 2).

Lexical choice varies with age. Therefore, interacting with people of different ages should lead to a more representative sample of naming patterns, leading to better prediction of people's lexical selection. Correct prediction is important as prediction is an integral aspect of language processing and is also related to comprehension ease and success. 55 participants were recruited via M-Turk. Participants reported how many people they regularly talk to in a week and the age range of their main group of interlocutors. Participants then performed a picture naming task on a set of pictures (N=33) for which name agreement norms for two age groups - college students and 60-75 year olds - exist. For each picture, participants provided the most common name that (1) college students and (2) 60-75 year olds would use to describe the picture. A mixed model analysis revealed that the wider the age range of participants' main social circle, the better they performed. So, having a more variable social circle makes one better at predicting people's lexical choice.

Additionally, the model tested whether having a larger social circle also improves people's lexical predictions (social circle size did not correlate with circle's age range). On the one hand, interacting with more people leads to a larger sample of naming patterns, and might therefore improve performance. On the other hand, the weight given to any token is reduced in larger samples, reducing the influence of interlocutors that are different from one's common interlocutor. Indeed, the same mixed model analysis revealed that participants with more interlocutors performed worse in predicting older adults', but not college students' responses. As most of our participants were young (M=33.5), and predicted better the responses of college students, these results show that a larger social circle can lead to worse performance with interlocutors different from one's common interaction partners.

Experiment 2 tested the influence of social circle size on global comprehension. Evaluative words (e.g., 'good') can express different degrees of positivity when used by different people. I tested participants' ability to decode the meaning that evaluative terms express when used by different people. Participants read elicited restaurant reviews from 2 reviewers: 6 reviews for each of 5 star levels for each reviewer (N=60). Participants estimated how many stars the reviewer assigned for the restaurant. The larger participants' social circle, the better they were at decoding the intensity of the review.

These studies show that the characteristics of one's social circle influence one's semantic skills. Specifically, having a more heterogeneous social circle boosts lexical prediction, while a larger social circle can boost comprehension of typical speakers but might be detrimental for performance with dissimilar interlocutors. This study opens up a window for studying how different aspects of one's social circle influence different linguistic skills.
TRAINING AND TIMING LOCAL ENRICHMENTS
Emmanuel Chemla (Ecole Normale Supérieure), Chris Cummins (University of Edinburgh), & Raj Singh (Carleton University)
c.r.cummins@gmail.com

Scalar implicatures (SIs) strengthen the meaning of a sentence with the negation of alternative sentences that could have been used. For example, “X is connected to some of its circles” conveys that X is connected to not all of its circles (the negation of “X is connected to all of its circles”). Such an SI is often found to come at a processing cost (e.g., Bott and Noveck, 2004).

Embedding, processing and external pressures. We investigated the behavior and processing profiles of scalar terms embedded in disjunctions. Taking some to be purely existential, the literal meaning of “X is connected to some or none of its circles” is tautologous; there is no way for it to be false. However, if some could be strengthened to some but not all within the scope of disjunction, the sentence would convey that “X is connected to only some or none of its circles”, and thus be false if (and only if) X is connected to all of its circles. The robustness of embedded strengthenings seems to be sensitive to contextual manipulations (e.g., Geurts & van Tiel 2013). Most relevantly, it has been argued that embedded strengthening of this kind can be driven by the need to satisfy independent constraints (e.g., Chierchia et al., 2012, Mayr and Romoli, 2014). In the above example, embedded strengthening – if available – would be motivated by the pressure to avoid tautologies.

Experiment. We tested the availability and processing consequences of this pressure for local enrichment using a self-paced reading paradigm. Participants read sentences of the form “X is connected to [Q] of its circles”, where X was a letter and Q a quantifier, and then evaluated whether a diagram presented on screen corresponded to that sentence. For the critical items (Q = some or none), two strategies are possible. S1: some is read with a local enrichment, yielding no tautology; or S2: some is first read literally and reanalysed after none is encountered, in order to avoid tautology. In the former case S1, some is predicted to be read slowly and “none” quickly, relative to baseline; in S2, the reverse is predicted to be true.

Training: local vs global. We trained two groups of participants (226 in all) on a distinct class of sentences (“Every letter is connected to some of its circles”). Training was achieved by providing feedback either compatible with local enrichments (False! if any one letter is connected to all of its circles – Local training condition) or with global enrichments (True! in the above case – Global training condition). After training, participants were exposed to 52 further sentences in pseudo-random order, including 6 critical none or some items.

Offline results. The tautology-avoidance pressure and the training were both effective. Embedded enrichments were high: displays in which X is connected to all of its circles were rejected when these followed some or none sentences (Global: 82%, Local: 92%), the Local group rejection rate being significantly higher (p < .01). The Local-trained participants were also significantly more reliable than the Global ones at enriching non-embedded some, as seen in their rejection of all displays following simple some sentences (95% v 83%, p < .01).

Online results. Participants in the Local training condition were, overall, slower readers than those in the Global condition (p < .01). Crucially, in the case of some or none sentences, Local participants were slower at the quantifier region but faster in the remainder of the sentence (as indexed by a significant interaction in a mixed regression model of reading times with full random effects structure: t = 2.39, p < .05). This suggests that the local enrichment spared these participants from needing to perform later reanalysis.

Conclusions. We conclude that (1) local enrichments are robust when there are pragmatic pressures that favour them, (2) these local enrichments, like global ones, are costly, and (3) local enrichments provide a mechanism to rescue otherwise deviant sentences as well as a processing advantage in such cases (no need for reanalysis). We discuss how these results speak to the relation between local and global enrichment procedures, which may or may not rely on the same mechanisms (e.g., grammatical vs neo-Gricean views), as well as the relation between enrichment and incremental interpretation strategies.
THE ROLE OF NUMERICAL COGNITION IN SIGNALLING QUESTION UNDER DISCUSSION

Chris Cummins (University of Edinburgh)
c.r.cummins@gmail.com

A recent debate in experimental pragmatics concerns the availability of quantity implicatures from expressions such as “more than n”, where n is a numeral. Fox and Hackl (2006) observe that “more than two” fails to convey “not more than three”, and argue on this basis that such expressions do not participate in implicatural scales. However, Cummins, Sauerland and Solt (2012) show that such implicatures are available from larger values of n, with “more than 80” reliably conveying “not more than 100”, although this pattern is less predictable when numbers are used that are not round (“more than 77”, etc.)

One possible explanation for this divergence in behaviour is that the use of round numbers is favoured because of their higher cognitive salience (Dehaene 1997, Butterworth 1999) and consequently the use of non-round numbers is taken to signal that these specific values are contextually important. For instance, (1) successfully conveys that 269 is a critical value for the topic under discussion, even if the reader does not know this in advance. By a similar token, the use of small cardinal values in expressions such as “more than three” could be taken to emphasise the potential relevance of those specific values (in this case, “three”).

(1) On this evidence, Obama is assured of more than 269 electoral college votes.

Theoretically, we could capture this by appeal to the idea of Question Under Discussion (QUD; Roberts 1996), and posit that the hearer of (1) infers that the QUD is, specifically, whether or not “more than 269” is the case, rather than simply “how many is the case”. For such a specific QUD, a stronger alternative such as “more than 300” would not necessarily achieve additional relevance, and therefore quantity implicatures such as “not more than 300” should be unavailable from (1) (cf. Breheny, Katsos and Williams 2006).

I investigate this claim by presenting participants with real-life usage examples of quantity expressions and asking them to judge (i) whether an implicature of the kind predicted by Cummins et al. (2012) is available and (ii) whether the specific number used was chosen for a particular reason. Participants were recruited via Amazon Mechanical Turk and responded to each question on a 5-point Likert scale. They were each asked to judge 12 examples drawn from the BNC, representing three types of numeral: large round, large non-round, and small integers. Across all conditions, there was a strong negative correlation between the mean judgments for (i) and (ii) (Pearson’s r = -0.67), indicating that participants refrain from drawing implicatures when they think the number has been chosen for a particular reason, a result that I interpret as consistent with a broadly Gricean account of implicature.

Comparing the three categories of number under test, participants rated sentences with non-round and small integers significantly above those with round numbers with respect to criterion (ii), and those with small integers significantly below the other categories on criterion (i) (Mann-Whitney U; all p < .05). This supports the idea that small integers tend to convey the existence of specific QUDs that do not license implicatures, thus potentially explaining the non-availability of the inference from “more than three” to “not more than four”. It also suggests that the use of large non-round numbers, such as 77, is an intermediate case, giving rise to implicatures and at the same time signalling the relevance of the specific number. I argue that this also makes sense within a Gricean framework, if we allow that the choice of numerical expression is informed by considerations of numerical salience.

I discuss the implications of this result for the question of how we compute QUDs, and to what extent they are helpful in understanding implicature. I also consider what this tells us about the usefulness of numerical expressions as a guide to how pragmatic enrichments are computed, and what pragmatics might tell us about the structure of the number system.
Focus particles like *only* access an alternative set [1]. For example, when hearing a sentence like *Anna only ate a banana*, alternatives to the word *banana* (such as *apple*) become salient to the listener [2]. Previous work has indicated that generating an alternative set involves activation and inhibition of multiple semantic associates [3]. Here, we investigate the impact of the particle *only* on the mechanisms involved in establishing alternative sets. In particular, we compare the retrieval of mentioned and unmentioned alternatives.

In Experiment 1, participants listened to short discourses (S1: *In the fruit bowl, there are pears, cherries and bananas. I bet Anna ate cherries and pears*). The manipulation was whether the focus particle *only* appeared in the third critical sentence or not (S2: *No, she _/only ate [bananas]_*). Subsequently, a target appeared on the screen that was either a mentioned alternative to the focused word (PEARS), an unmentioned alternative of the same semantic category (APPLES) or an unrelated word (SOCKS). Participants had to indicate whether the word had been mentioned or not (probe recognition). The results showed that the recognition of mentioned alternatives was slowest, the rejection of unmentioned alternatives intermediate and the rejection of unrelated items fastest. The particle *only* slowed reactions relative to the bare condition for mentioned and unmentioned alternatives but not unrelated items. The selective interference effect of *only* suggests that identifying the relevant alternatives is subject to competition among semantic associates. This competition might be stronger in the case of *only* compared to bare focus because *only* establishes a conventional association between the focused element and its alternatives [1]. The overall differences across target types suggest that a comparison is being made between members of the alternative set in a probe recognition task. We surmise that the focused element (e.g., *bananas* in the example) takes part in this competition, decreasing participants’ ability to match the mentioned alternatives with the prior discourse.

Experiment 2 further explored the underlying mechanisms. We used the same materials and basic procedure but participants were required to indicate whether a word existed or not (lexical decision). The overall pattern of results was reversed: Reaction times were fastest for mentioned alternatives, intermediate for unmentioned alternatives and slowest for unrelated items. These results indicate that mentioned alternatives receive the highest activation and that additional unmentioned alternatives become activated through semantic priming. Additionally, there was an overall interference effect of the particle *only* relative to the bare condition, in line with the competition hypothesis.

Overall our experiments show that listeners entertain a set of mentioned as well as unmentioned alternatives when processing focal information (even when the context is restricted to a set of element). The results further indicate that alternative sets are established by competitive inhibition among the semantic cohort. We assume that, initially, a cohort of semantic alternatives is accessed from the mental lexicon and that competitive inhibition renders the relevant alternatives salient (see also [3] for a similar conclusion).

EMBODIED CONCEPTUALIZATION IN LANGUAGE PRODUCTION

In language production speakers come up with a message that encodes aspects of the apprehended event (conceptualization) before mapping the message onto syntax and phonology. Empirical research into how people conceptualize for speaking has been scarce, which may reflect the fact that theories of conceptualization based on amodal, ungrounded symbols are still prevalent in production research, in contrast to comprehension where embodiment has provided a novel framework for research¹. In this paper, capitalizing on recent findings that conceptual processing in language comprehension and other cognitive domains is subserved by sensorimotor simulation², we investigated whether conceptualization in language production also makes use of similar sensorimotor simulations. To do this, we devised a novel mouse click paradigm in which participants make clicks horizontally across the screen while describing an event (one click per word) (Fig. 1A). In Experiment 1, 21 participants described 32 pictures depicting dative events (e.g., professor giving sailor banana) from the perspective of the boxed character, which was either the agent (e.g., I give the swimmer a banana) or the recipient (e.g., The professor gives me a banana). If speakers mentally simulate the event, the event should be simulated as an outward transfer of the banana from the perspective of the agent (professor) but as an inward transfer from the perspective of the recipient (sailor). These mental simulations should therefore influence concurrent mouse clicks such that the outward/inward mental simulation should respectively bias the mouse to move outward (upward onscreen) or inward (downward onscreen). LME modelling of the results supported this hypothesis: participants made more downward trajectories when they described pictures from the recipient’s perspective than from the agent’s perspective (Fig. 1B). These findings suggest that speakers use sensorimotor simulation egocentrically (i.e., projecting an event-participant onto themselves to mentally reenact an event) when conceptualize for first-person descriptions. Experiment 2 tested whether speakers also employ similar egocentric mental simulation for third-person descriptions. 45 participants began the picture description with the boxed character (e.g., The professor gives the swimmer a banana or The swimmer is given a banana by the professor). We found no difference in the trajectories (Fig. 1C) whether the description began with the agent or the recipient. In all, the study presents the first evidence that speakers use sensorimotor simulation to conceptualize for speaking. We discuss how embodied conceptualization may explain previous findings of conceptual effects on grammatical encoding.

Fig. 1: The experiment procedure (A), results of Experiment 1 (B) and Experiment 2 (C).

INSTRUMENT PRIMING: A TALE OF TWO STRATEGIES
Chung-I Erica Su (Academia Sinica) & Gail Mauner (University at Buffalo)
ericasuci@gmail.com

Priming is often used to explore conceptual organization. For instance, Ferretti et al. (2001: Exp't 2) used instrument/location related/unrelated verb-noun pairs (e.g., stirred-spoon, shot-spoon; cooked-kitchen, worshipped-kitchen) in a semantic priming task to test the hypothesis that the arguments of verbs are verb-specific concepts. They observed significant priming only for instruments, and assumed it was the consequence of verbs automatically activating prototypical instrument role fillers because a short stimulus onset asynchrony (SOA) and small relatedness proportion (RP) were used. By comparing the magnitudes of instrument priming across experiments with unrelated verb primes and psychological verbs (e.g., urged-spoon; awakened-kitchen) as conceptually neutral primes, we examined whether instrument priming was automatic, or instead, due to two strategies.

Experiment 1, which was an exact replication of Ferretti et al.’s Exp’t 2, yielded significant instrument priming that accounted for 16% of the variance (see Table 1). Our Exp’t 2 replaced unrelated instrument/location verbs with neutral verbs as primes. Since related verb-noun pairs were identical in Exp’ts. 1-2, magnitudes of instrument priming should be similar if verbs automatically prime prototypical instrument role fillers. Instead, instrument priming was not significant and accounted for only 5% of the variance. This indicates that little if any of the instrument priming was due to automatic facilitation.

We assume that strategy formation exploits salient regularities and that participants are biased to adopt more successful strategies before less successful strategies. A sorting task revealed that location targets were salient since locations were grouped into fewer coherent categories (M=3.6) than instruments (M=8.9). We hypothesized that the salience of location targets led to the formation of a location anticipation strategy, which successfully predicts targets on 50% of experimental trials. A second strategy, predicting instrument targets following verbs that describe events typically involving instruments, successfully predicts instruments on 25% of these trials. Use of these two strategies would explain null location and significant instrument priming in Exp’t. 1.

Exp’ts. 3-4 examined whether the instrument and location priming results in Exp’t. 1 were due to an instrument expectation strategy whose development was facilitated by adoption of a location expectation strategy. Exp’ts. 3-4 were replications of Exp’ts. 1-2, except that location prime-target pairs were omitted. The magnitude of instrument priming was reduced in Exp’t 3, accounting for only 9% of the variance (see Table 1). In Exp’t 4, when related instrument verb primes were replaced with neutral verbs, instrument priming accounted for only 1% of the variance. Thus, the absence of a location expectation strategy impeded the formation of an instrument expectation strategy. Thus, the instrument priming in Exp’t. 1 was due to violations of instrument target expectations following unrelated instrument verb primes. Our results show that predictive strategies develop even when “best practice” 250 ms SOA and 0.25 RP thresholds are employed, and that the ease or odds of formulating a strategy may depend on whether an easier or more successful strategy is already in use.

Table 1: Lexical decision times (ms), priming effects (Unrelated/Neutral-Related differences) and effect sizes ($\eta^2$) for Experiments 1-4 (* indicate significance at $p < .05)$.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Location Baseline</th>
<th>Related</th>
<th>Priming</th>
<th>Instrument Baseline</th>
<th>Related</th>
<th>Priming</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Unrelated verbs</td>
<td>809</td>
<td>812</td>
<td>-3</td>
<td>823</td>
<td>781</td>
<td>42*</td>
<td>.16</td>
</tr>
<tr>
<td>2 Neutral verbs</td>
<td>785</td>
<td>788</td>
<td>-3</td>
<td>807</td>
<td>784</td>
<td>23</td>
<td>.05</td>
</tr>
<tr>
<td>3 Unrelated verbs</td>
<td>772</td>
<td>743</td>
<td>29*</td>
<td>777</td>
<td>9</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>4 Neutral verbs</td>
<td>786</td>
<td>777</td>
<td>9</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CATEGORIZATION AND IMPLICIT LEARNING OF SEMANTIC ROLES
Eva Wittenberg, Manizeh Khan, & Jesse Snedeker (Harvard University)
eva@wjh.harvard.edu

How do we get from a syntactic structure to an interpretation? While often, there is a one-to-one mapping between syntax and semantics, so-called light-verb constructions (LVCs) like "Julius gave a kiss to Anne", in which the action (kissing) is denoted by a noun in a syntactic argument position, pose a problem for such a direct mapping. Recent research suggests that processing LVCs is associated with a different neural response than processing non-light sentences such as "Julius gave a toy to Anne" ([1]). A behavioral study has found that people do not consistently sort Agent-Patient events described by LVCs with their base-verb Agent-Patient counterparts ("Julius gave a kiss to Anne"), but sometimes with their non-light Source-Goal-Theme counterparts ("Julius gave a toy to Anne"; [2]). How do they arrive at this intermediate interpretation? Either, both the Agent-Patient and the Source-Goal-Theme mapping are considered at the same time ("argument sharing", [3]); or speakers first consider only the canonical Source-Goal-Theme mapping, and then shift imperfectly to an Agent-Patient representation. We used an eye-tracking study to investigate these hypotheses.

Participants implicitly learned to categorize sentences according to number of semantic roles. Following the methodology in [4], we presented sentences auditorily while a ball rolled into a Y-shaped tube; participants’ task was to click on it as soon as it emerged. On 25% of trials, the ball was “stuck”, and participants had to guess where it would have reappeared by clicking on one side or the other. Mouse clicks (Figure 1) show that they learned to associate two-role sentences (“John ate an apple”) with one side of the screen, and three-role sentences (“Jan stole a handbag from the old lady”, “Bert handed the teacher his notebook”) with the other side. Eye-movements (Figure 2; analyzed with a non-parametric permutation test [5]) confirm this, with looks for two-and three-role sentences diverging from each other early on (t=85.6, p<.01, 600-2900ms). Looks for LVCs, however, were intermediate and only diverged from both two-role and three-role sentences after the sentence ended (two-role vs. LVC: t=30.4, p<.01, 1600-2800ms; three-role vs. LVC: t=6.8, p<.01, 1800-2100ms; t=14.9, p<.01, 2400-2900ms).

These results suggest the comprehender considers two mappings in parallel ("argument sharing", [3]). We discuss these findings from a theoretical perspective as well as in conjunction with recent behavioral and on-line data ([1],[2]).

Global ambiguity of modifier attachment to one of the nouns in a complex noun phrase (e.g. the servant of the actress that was on the balcony) was supposed to be resolved according to the Late Closure Principle (Frazier & Fodor 1978), which favours low attachment (LA). However, after high attachment (HA) preference was discovered in Spanish, contradicting the idea of universality of parsing principles (Cuetos & Mitchell 1988), dozens of cross-linguistic studies have been carried out, classifying languages into two groups: favouring LA or HA, and trying to explain the difference. Data reported for Russian relative clauses (RCs) are controversial (Sekerina 2003, Fedorova & Yanovich 2004). Studying Russian participial constructions with HA, LA and ambiguous attachment may shed some light on this problem.

**Method and participants.** 60 native speakers of Russian performed a word-by-word self-paced reading task on Presentation software. **Materials and design.** 24 sets of target stimuli, as in (1a-c), were constructed. In each sentence a complex noun phrase was followed by a participial construction, which could be attached either to the first or to the second NP. The case form of the participial either disambiguated the modifier towards N1 or N2 or left it ambiguous. Both interpretations of the ambiguous sentences were judged as equally plausible by 32 native speakers. N1 and N2 always had the same number and gender, animacy was balanced across sets. All participial constructions had roughly the same length (12-13 syllables). Every participant saw each target sentence once, in one of the conditions. All sentence (including 32 fillers) were followed by a question with two possible answers (e.g. ‘Who saw the robbery? The driver, the workmate’ in the set (1a-c)).

(1) a. Svidetel’ upomjanul naparnika voditelja, včera videvšego eto ograblenie. Cond. AMB. Witness mention workmateACC driverGEN yesterday having-seenACC-GEN this robbery

   b. Svidetel’ upomjanul o naparnike voditelja, včera videvšego eto ograblenie. Cond. LA. Witness mention about workmatePREP driverGEN yesterday having-seenGEN this robbery

   c. Svidetel’ upomjanul o naparnike voditelja, včera videvšem eto ograblenie. Cond. HA. Witness mention about workmatePREP driverGEN yesterday having-seenPREP this robbery

**Results.** The analysis of reading times by RM ANOVA shows an overall preference for LA: sentences disambiguated towards LA are read faster than sentences disambiguated towards HA and ambiguous sentences in the participle region (significant in subject and item analyses) and the following region (significant in subject analyses). Moreover, if RTs for AMB sentences are divided depending on the answer a participant eventually chooses, we see that sentences with HA answers take longer than sentences with LA answers. We can conclude that (1) LA is easier to process and (2) ambiguity has a processing cost. But the answer analysis for AMB sentences shows a 60.8% dominance of HA interpretations (significant according to the χ² test). Thus, online and offline preferences do not coincide.

**Discussion.** We argue that in accordance with the Late Closure principle, LA is more economical. Why speakers nevertheless do not tend to choose the easiest option can be explained by Fodor’s (1998) prosodic principle: Russian words are often long, so RCs and participial constructions are usually heavy.


---

1 The study was supported by the Russian Foundation for Fundamental Research (grant #12-06-00382-a).
Native listeners are sensitive to deviations from typical native productions and the degree of deviation affects the perception of foreign accentedness. However, the gradient nature of foreign accentedness has not yet been thoroughly examined in lexical processing. Here we investigate the role of gradient foreign accentedness in the activation of lexical representations and in lexical retrieval, using cross-modal identity priming and visual word eye-tracking, respectively. Assuming representations contain detailed information (cf. Goldinger, 1998), we expect activation and access to vary with gradient foreign accentedness and by listener experience with foreign-accented speech.

Forty monosyllabic English words were extracted from the Wildcat corpus (Van Engen, et al., 2010), recorded from four talkers (three native Chinese, one native English). Serving as auditory tokens, they represent the full scale of accentedness ratings (Porretta & Tucker, 2012). For the priming experiment, each was matched with a written target: identity, unrelated word, or pseudo-word. Responses and reaction times of 48 participants were recorded during lexical decision. For the eye-tracking experiment, each token was paired with four written words: identity, onset competitor, rhyme competitor, and distractor (cf. Allopenna, et al., 1998). Eye movements were recorded from another group of 48 participants while locating the written form of the token. All participants responded to a questionnaire assessing their experience interacting with Chinese-accented speakers.

Generalized additive mixed modeling (GAMM) of the reaction times from the priming study reveals that, unsurprisingly, reaction times are significantly faster when the target word matched the auditory prime (Wood, 2006). Interestingly, however, when examining only these Prime-Identity trials, differential non-linear effects of accentedness rating emerge between experience groups. Namely, high experience listeners have lower reaction times overall which then spike for very highly accented primes. Low experience listeners have higher reaction times, displaying a gradient increase across the accentedness continuum.

In the eye-tracking study, logit-transformed fixations to target were modeled over the period of 200–700 ms post stimulus onset using GAMM. Overall, increased accentedness results in decreased fixations, with a non-linear cline across the continuum. High experience listeners show increasing fixations over time which attenuate for highly accented items. Low experience listeners show a similar, yet flatter surface over time and accentedness.

The results indicate that native-like productions facilitate lexical access more than productions that deviate from native norms. The priming study suggests that native and weak accent produce stronger lexical activation, while the eye-tracking study suggests that these tokens result in more rapid access. In both, this facilitation is modulated by the listener’s experience with Chinese-accented speakers. Thus, greater experience may lead to more detailed representations on which to map variability. This may indicate that, even in adulthood, lexical representations are shaped through experience with variation.

References
LISTENERS RECOGNIZE OTHERS' SPEECH BETTER THAN THEIR OWN
Will Schuerman (MPI for Psycholinguistics), Antje Meyer (MPI for Psycholinguistics), & James McQueen (Donders Centre for Cognition, MPI for Psycholinguistics).
Will.Schuerman@mpi.nl

To what extent does speech perception depend on our own experience in speech production? One idea is that we find our own speech more intelligible than that of others. Experiments on action perception (e.g. dart throwing [1], handwriting [2], and lip-reading [3]) found increased accuracy if stimuli had been produced by the same participant as compared to another. This would suggest that our comprehension is facilitated when a perceived action resembles the representations accessed when we produce that same action. An alternative view holds that incoming speech is decoded by reference to a statistical average of our linguistic input, rather than by our own history of production. This would predict that our phonetic systems are more attuned to the input coming from a prototypical speaker of our linguistic community, not our own personal idiosyncratic productions.

To investigate this, we asked speakers to identify words produced by themselves vs. another speaker. The experiment consisted of two phases: a Production phase and an Identification phase. In the Production phase, 28 female native speakers of Dutch read aloud 120 Dutch words. Words were equally divided into "Easy" and "Hard" Conditions according to frequency, phonological neighborhood density and average neighborhood frequency. The duration, amplitude, and other phonetic parameters of recordings were compared using principal component analysis. The participant with the smallest average distance to all other speakers was chosen as the 'Model' speaker, leaving 27 participants for the Identification phase.

In the Identification phase, participants listened to degraded (6-band noise vocoded) versions of the recorded words and tried to identify each word by typing in their response. Noise-vocoding removes many cues associated with speaker identity while leaving durational and amplitude information intact. Half of the degraded stimuli were drawn from the participant's own recordings (Talker = "Self") while the other half were drawn from the recordings of the Model speaker (Talker = "Other"). Participant responses were transcribed phonemically to enable calculation of Levenshtein distances between a target word and the participant's response (0 = correct). Average Levenshtein Distances are reported in Table 1.

Words in the Easy List were more accurately identified than Hard words; participants were on average more accurate for Other produced words than Self produced words. Data was analyzed using a hurdle model, which is appropriate for non-normally distributed data with a high number of zero responses. This is in effect two models; a logistic regression model on data coded as zero and non-zero, and a zero-truncated poisson model to examine Levenshtein distances greater than 0. In both models, significant main effects were found for Word Difficulty as well as Talker.

The results suggest that speakers are more accurate at identifying speech that has been produced by a statistically average speaker than by themselves. This is consistent with the view that the phonetic representations accessed during perception are more reflective of the overall input of our speech community rather than our experience with our own idiosyncratic productions. While not denying an integral role for production experience in perception, we argue for a more nuanced view of speech perception that takes into account the demands of understanding a wide range of talkers.

PHONOLOGICAL REPRESENTATIONS IN ADULTS WITH DYSLEXIA: EVIDENCE FROM A PERCEPTUAL LEARNING PARADIGM
Margriet A. Groen (Behavioural Science Institute, Radboud University Nijmegen, Netherlands) & James M. McQueen (Behavioural Science Institute & Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands).
m.groen@pwo.ru.nl

Deficient phonological representations have been put forward as a hallmark of developmental dyslexia. One proposal is that dyslexics' representations of speech sounds are not sufficiently specified (e.g., Manis et al., 1997; Mody et al. 1997). Others have suggested that the representations are fully specified, but that dyslexics fail at tasks that are demanding in terms of access to those representations (Boets et al., 2013; Hulme & Snowling, 1992; Ramus & Szenkovits, 2008; Shankweiler & Crain, 1986).

When listening to speech, we need to adjust our interpretations of speech cues in response to talker-specific differences in articulation (Ladefoged, 1989; Ladefoged & Broadbent, 1957). The task of dealing with talker variability is usually handled with ease by the human perceptual system and listeners use lexical knowledge to retune speech-sound perception (Norris et al., 2003). Exposure to an ambiguous sound [?], that was midway between [f] and [s], caused a shift of the [f]-[s] category boundary when [?] was placed in contexts that were lexically consistent with its interpretation as either [f] or [s]. Subsequent work (McQueen et al. 2006) showing generalisation of learning suggests that this retuning involves phonologically abstract representations. This offers the opportunity to assess the quality of and access to abstract phonological representations without relying on reading in people with dyslexia.

We employed such a perceptual-learning paradigm with speech sounds in combination with an implicit lexical priming task—tapping representation—and explicit phoneme identification and phoneme categorisation tasks—tapping access. To expose participants to the ambiguous sound [?] in a lexical context, we used a story (Eisner & McQueen, 2006) in which either all [f] or all [s] sounds were replaced with an ambiguous sound [?] midway between [f] and [s]. Subsequently, participants completed an auditory version of an identity priming task ('implicit' task; McQueen et al., 2006) and one of two explicit tasks: a phoneme discrimination (AXB) task or a phoneme identification task.

Participants were 42 adults with dyslexia and 48 controls. Adults with dyslexia had received a clinical diagnosis of developmental dyslexia as a child and scored below the 24th percentile on standardised measures of word and non-word reading. Controls had no history of reading problems and scored well above this cut-off. Additionally, adults with dyslexia performed more poorly on measures of phonological awareness (spoonerisms) and lexical retrieval (rapid naming), but performed similarly, in terms of response pattern, to controls on the implicit and explicit experimental tasks. In particular, dyslexics and controls appeared equally able to use knowledge about how spoken words ought to sound to learn how to interpret ambiguous speech sounds. Nevertheless, adults with dyslexia were slower in responding on the lexical priming and the phoneme discrimination tasks.

These results indicate that the phonological representations of dyslexics might be both fully specified and accessible, although access might be slower.

OVERSPECIFICATION IN REFERENCE: MODELLING SIZE CONTRAST EFFECTS

Roger P.G. van Gompel (University of Dundee), Albert Gatt (University of Malta), Emiel Krahmer (University of Tilburg), Kees van Deemter (University of Aberdeen)
r.p.g.vangompel@dundee.ac.uk

Speakers often overspecify: They provide more information than necessary to identify the target (e.g., Englehardt et al., 2006; Deutsch & Pechmann, 1982). It is often assumed that overspecification is due to the order in which properties are selected. According to the Incremental Algorithm (IA, Dale & Reiter, 1995), speakers first select the most salient property (colour) for the circled target in Fig. 1a and subsequently add size to rule out the remaining distractor, but this makes colour redundant and results in the overspecified small black candle. In contrast, in Figs. 1b and 1c, speakers also first select colour, but because this rules out all distractors, they do not add size, resulting in black candle.

Colour may be intrinsically more salient than size, because size is a relative property that requires comparison with the distractors (Pechmann, 1989) or because it is more central in conceptual gestalts than colour (Belke, 2006). However, the colour-over-size preference in previous studies may have occurred because the colour manipulation was stronger than the size manipulation. If the size contrast is larger (Fig. 2a-c), colour may no longer be preferred. This may affect overspecification: According to IA, if size is selected before colour, speakers produce overspecifications in Fig. 2b, but not in Fig. 2a. (No overspecifications should be produced in Figs 1c, 2c, because the initially selected property is always fully distinguishing.)

To test this, participants described the targets in Figs. 1-2 to another participant, who had to select the target. The data (below the Figs.) showed that with a large size contrast, colour was no longer preferred over size when either property was fully distinguishing (Fig. 2c) and colour alone was used no more often in Fig. 2b than size alone in Fig. 2a.

To test IA, we modified it so the first property (colour or size) was chosen probabilistically (contra Dale & Reiter, 1995), with the probability determined by the colour-size preference. To predict how often speakers used each expression in a condition, we used the colour-size preference determined on the basis of the two other conditions with the same size contrast (e.g., for Fig. 1a predictions, we used Figs. 1b, 1c; for Fig. 2b, we used Figs. 2a, 2c).

Because probabilistic IA does not account for overspecification in Figs. 1c and 2c (predictions below the Figs.), we tested the probabilistic overspecification model (PRO, Gatt et al., 2012). It assumes that speakers always first select the property that is fully distinguishing (size in Figs. 1a, 2a; colour in Figs. 1b, 1c). If two properties are fully distinguishing (Figs. 1c, 2c), then the first property is probabilistically selected according to preference. Next, speakers add the second property, with the probability of adding it depending on its preference. As before, the predictions for each condition were based on the best-fitting parameter values for the two other conditions with the same size contrast. The PRO predictions (shown below the Figs.) provided a good fit with the data.

The experiment shows that the colour-over-size preference in previous studies occurred because the size manipulation was not strong enough and that the magnitude of the size contrast strongly affects whether overspecification occurs. The modelling shows that PRO provides a better fit than probabilistic IA, suggesting that overspecification does not occur because an initially selected property becomes redundant (as in IA), but because speakers often add a property even if the initially selected property is fully distinguishing (PRO).
AN INFORMATION-THEORETIC ACCOUNT OF WORD MEMORABILITY
Kyle Mahowald, Phillip Isola, Evelina Fedorenko, Aude Oliva, Edward Gibson (MIT)
kylemaho@mit.edu

Building on existing work on verbal memory [e.g., Shepard, 1967; Wickelgren, 1972] and advances in visual memory [e.g., Isola et al., 2011], we ask what makes a word memorable and develop a Bayesian model of how an optimal agent with finite resources might store and retrieve words. Estimates derived for each word using this rational analysis account explain a significant amount of the between-word variance in human performance.

To measure the memorability of 2,222 words (sampled from the Subtlex database of movie subtitles and hand filtered to remove morphologically related words and alternate spellings), we used a repeat detection task in which participants (n=676) played a “Memory Game” online in which they were asked to detect repeats in a long stream of words.

People’s ability to remember words was on par with memory for scenes and better than memory for faces in similar paradigms [Isola et al., 2011; Bainbridge et al., 2013]: mean hit rate .68; mean false alarm rate .10; accuracy .80 [where accuracy is (correct hits + correct rejections)/(total presentations)]. Although accuracy was high, some words were consistently better remembered than others (split-half Spearman correlation across participants: .58). To explain the consistent variation across words, we propose a Bayesian model of word recognition memory. Our model predicts that recognition accuracy will be related to an information-theoretic metric \( M \), formalized as: \( M(w) = H(r|w) - E_P(r|w)[\log(P(r))] \), where \( r \) is the intended referent and \( w \) is the specific word used. This measure is derived by assuming that, during the task, when a participant sees word \( w \) she stores not the word but a referent \( r \) selected by that word. When the participant sees a new word, she decides how likely it is that this same word could have generated the remembered referent \( r \).

\( M \) decomposes into two terms: the entropy of the conditional distribution of the referent given the word (related to how many meanings the word has) and a measure of how unique a given word is in its ability to refer to \( r \) (related to how many synonyms a word has).

Thus, we predict that participants will better recall words with few synonyms and meanings. A word like last is ambiguous in meaning in that it could be an adjective describing order, or a verb (“keep going”). Each sense also has many synonyms: ultimate, final, concluding, etc. for the adjective. A term like Bob Marley, however, refers to one entity and has no synonyms.

According to our Bayesian model, word memorability can be predicted from three terms: number of synonyms, number of meanings (these can be estimated from Wordnet or behavioral ratings on Mechanical Turk; both estimates produce similar results), and frequency. All 3 are strongly correlated with task accuracy in the directions predicted by the model (Spearman \( r \) of .54, .29, .43, respectively), and we can combine them linearly, using simple regression, to approximate \( M \). Training and testing on entirely different words and participants (split-half) to avoid overfitting, the mean Spearman correlation (over 1000 iterations) between the regression’s predictions and the out-of-sample empirical values was .49 (lower than some of the in-sample correlations above due to less data being available in cross-validation). Including other ratings from MTurk (arousal, concreteness, familiarity, image ability, and valence) in the model, the correlation is .56 (almost as high as the split-half inter-subject consistency of .58, suggesting the model explains most of the explainable variance across words).

These results suggest that people use memory resources efficiently. Moreover, the account presented here may be used to improve performance in NLP applications involving text memorability (such as advertising and education), where frequency alone is used as a proxy for lexical memorability (Danescu, et al. 2012).
COMMON GROUND AND JOINT UTTERANCE PRODUCTION: EVIDENCE FROM THE WORD CHAIN TASK

Jarosław Lelonkiewicz (University of Edinburgh), Chiara Gambi (Saarland University)
J.R.Lelonkiewicz@sms.ed.ac.uk

Instances in which one interlocutor continues an utterance initiated by another are not infrequent in conversation (Howes et al., 2011). Howes et al. (2012) found that participants were more likely to continue an artificially truncated turn if it was about the current conversation topic. It is possible that this preference is related to interlocutors finding it easier to predict one another’s utterances when these utterances are about something in common ground (see Pickering & Garrod, 2013).

In our study, we manipulated common ground by using two sets of 20 ambiguous (M\text{CELEX frequency} = 1107 p.m.; M\text{length} = 1.4 syllables) and 20 non-ambiguous words (M\text{CELEX frequency} = 1211 p.m.; M\text{length} = 1.4 syllables; p’s > .2), for which 18 pairs of participants were asked to produce definitions. The ambiguous words were balanced (dominant meaning frequency ≤ .65 and ≥ .41). We hypothesised that joint production would proceed more smoothly if participants were able to assume shared meanings with their partner (as should be the case with non-ambiguous words), because this would constrain their predictions about what will be uttered next.

Participants were seated in separate booths and interacted via a text-based chat environment (DIET chat-tool; http://cogsci.eecs.qmul.ac.uk/diet/). They were presented with one word at a time and had to jointly construct a definition for each word. Instructions emphasized speed but also that the definition had to provide sufficient information to allow a third party to guess the word. As in the American TV game Chain Reaction, the participants could contribute only one word per turn, and had to continuously switch turns with their partner (see 1, produced as a definition of BAT). Although natural joint production lacks such a constraint, it similarly requires incremental interpretation and tight yoking of comprehension and production processes.


As a control, 26 participants provided definitions for the same words in a solo version of our task. Similarly to those working together, solo participants could type only one word per turn, but were working entirely on their own. We measured the total time spent typing and the number of words produced per definition, and computed typing speed as number of words per second. Typing speed was higher for non-ambiguous than ambiguous words when participants were interacting with another (M\text{non-amb} = .50, M\text{amb} = .46), but not in the solo task (M\text{non-amb} = .71, M\text{amb} = .71; Ambiguity X Task interaction: \( \chi^2 (1) = 4.21, p < .05; \) maximal random effects structure).

In sum, we showed that jointly producing an utterance is more difficult when common ground cannot be assumed but needs to be established. Additional analyses will confirm whether typing speed is affected predominantly at the beginning of definitions for both ambiguous and non-ambiguous items, reflecting the cost of establishing common ground.

How do people label objects in dialogue, and does this affect how they conceptualise them?

For the collaborative account (Brennan & Clark, 1996) choice of labels is contextual and driven by agreement – a ‘conceptual pact’ – between interlocutors on a shared label, which reflects a common conceptualisation of the referent entering their ‘common ground’. For the competing alignment account, choice of labels is instead driven by implicit repetition of recently processed labels, perhaps through forward modelling (Pickering & Garrod, 2013). The alignment account predicts that when interlocutors use the same labels, they will come to privately conceptualise their referents in the same way too: but if so, this conceptualisation process could admit interlocutor-internal interference from each interlocutor’s other concepts, and perhaps project such interference to these concepts as well – linking label choice in dialogue to conceptual revision.

In this study, groups of three participants played a tangram-matching game (Clark & Wilkes-Gibbs, 1986) in successively interacting pairs, with the Matcher in a round becoming Director for another player in the next round – an arrangement similar to Garrod & Doherty (1994). We manipulated whether Player C (acting as Matcher in round 2 and Director in round 3) saw and pre-conceptualised (by privately conceiving labels for them) either the experimental tangrams or a foil set of tangrams before the game began. Player C thus played the game either having previously conceptualised and conceived possible labels for the experimental tangrams or not.

Overall, participants created stable labels (e.g. “the kung fu man”) that carried across interactions from round 1 to round 2 despite the changing pairings: when acting as Matcher in round 2 Player C therefore accepted the labels used in the game so far (“kung fu man”) over repeated exchanges with Player B. However, in round 3 there was a significant likelihood (analysed by multilevel logistic regression) that if Player C had privately labelled the tangrams before the game they would abruptly switch to different labels (p < .05): specifically those labels they had originally conceived (e.g. “the sinking boat”). These results cannot be explained using the collaborative model, as participants explicitly violated the existing, contextually-established pact (that participants in the foil-labelling condition still maintained in round 3). But nor can they be explained in terms of simple alignment, as participants had repeatedly processed the other labels previously, only to then stop using them. These results imply that conceptual interference internal to the speaker may occur during dialogue more than currently envisioned by alignment or a purely contextual account of label choice. Moreover, preliminary data indicates the inverse may be true as well: label choices in a dialogue may immediately propagate back to and revise prior conceptualisations of their referent, arguing against a simple ‘director role effect’ account.
THE RELATIONSHIP BETWEEN IMPLICIT EXPECTATIONS ABOUT CHARACTER BEHAVIOUR AND EATING DISORDER TENDENCIES: EVIDENCE FROM EYE MOVEMENTS DURING READING

Christina Ralph-Nearman (University of Nottingham) & Ruth Filik (University of Nottingham)
Lpxcr4@nottingham.ac.uk

Eating disorders have a higher mortality rate than all other psychiatric disorders combined (NICE, 2004). There are several theories surrounding which psychological factors are associated with the development and maintenance of an eating disorder. Firstly, there is the Integrated Cognitive-Behavioral theory (Williamson, White, York-Crowe, & Stewart, 2004), in which body self-schema is a primary cognitive concept, with selective attention to disliked body parts and body shape being theorized as key in the etiology of eating disorders. Secondly, there is the hypothesis that those with eating disorder tendencies have food-related cognitive biases, specifically, an attentional bias toward negative food-related and eating-related stimuli (Shafran, Lee, Cooper, Palmer, & Fairburn, 2007). Finally, perfectionism is also often cited as a key aspect in many different eating disorder theories, such as the Two-Factor Theory (Joiner, Heatherton, & Keel, 1997), the Three-Factor Theory (Bardone-Cone, Joiner, Crosby, Crow, Klein, le Grange, et al., 2008), and Fairburn, Cooper, and Shafran’s (2003) Transdiagnostic Model of Eating Disorders. Our aim was to use eye-tracking during reading to investigate whether participants’ implicit expectations regarding how characters will behave in body image-related, food-related, or perfectionism-related scenarios will be related to their tendency towards disordered eating behaviour.

Ninety native English-speaking females aged 18 to 38 had their eye movements monitored while they read 36 body-, food-, or perfectionism-related scenarios (e.g., 1-3), which ended with a target sentence containing a critical emotion-based word that either ‘matched’ or ‘mismatched’ with one’s expectations concerning how the character might react. All materials were pre-tested, to ensure that a) each scenario solely fit one condition (e.g., was either body, food, or perfectionism-related), and b) was viewed as either matching or mismatching one’s expectations. After the eye-tracking task, participants completed the Eating Disorder Examination Questionnaire (EDE-Q 6.0; Fairburn & Beglin, 2008).

1. Body image-related scenario: Jane walked in the room and chatted with a few people. Then, someone walked up and told her that she looked [slim /fat mismatch] that day.
   /She was pre-critical/ delighted critical/ to be told that. post-critical/
2. Food-related scenario: At her friend’s house, Jane ate a piece of cake. She was later told that the piece of cake had contained [very few /a thousand mismatch] calories.
   /She was pre-critical/ delighted critical/ to be told that. post-critical/
3. Perfectionism-related scenario: Jane worked hard on her assignment. When the marks came back, she was told that she had made [no /many mismatch] mistakes in her essay.
   /She was pre-critical/ delighted critical/ to be told that. post-critical/

The eye-tracking data showed no significant effects in first-pass reading times, but there were significantly longer total reading times for mismatching than matching conditions (i.e., a significant mismatch effect) in the pre-critical, critical, and post-critical regions, and this did not interact with scenario type (e.g., body vs. food vs. perfectionism). Results also showed that the size of this mismatch effect for perfectionism-related materials in total reading times for the pre-critical region was positively related to participants’ scores on the EDE-Q 6.0. There was no relationship between EDE-Q 6.0 score and the size of the mismatch effect observed in total reading times for food-related or body-related items. These findings support for theories which propose that perfectionism may be a key aspect of developing and maintaining an eating disorder such as the Two-Factor Theory, Three-Factor Theory, and Transdiagnostic Model of Eating Disorders. The observed relationship between personality factors and reading behaviour also highlights the need to consider inter-individual differences when developing models of on-line language processing (e.g., van den Brink et al., 2012).
The effect of bilingualism on cognitive processes has received much attention in recent literature. This study will contribute to the literature by examining Age of Immersion (AoI) effects on linguistic and non-linguistic processing in bilingual young adults. Bilinguals have been shown to have improved cognitive control skills (e.g. the abilities to pay attention to relevant input, to inhibit inappropriate responses, and to switch between tasks effectively) in linguistic tasks such as the Stroop test\(^1\) as well as in non-linguistic tasks such as the Attention Network Test\(^2\). These bilingual advantages may stem from the need to manage two simultaneously-active languages in one mind, thus requiring the recruitment of a domain-general control mechanism to mediate the conflict between the two languages. Much of the literature, however, has examined highly proficient bilinguals, with minimal attention to the age at which they were immersed in their second language. Furthermore, studies that do examine age of acquisition effects often collapse simultaneous and early sequential bilinguals into a single group, an approach which fails to capture any differences between these two potentially distinct groups.

The current study divided 40 English-French bilinguals into three AoI groups: Functional Monolinguals (n=14), who, while exposed to French, had never been immersed in the language; Simultaneous bilinguals (n=11) who were immersed in both languages from birth; and Early bilinguals (n=15), who began a French immersion programme before the age of 7. Participants took part in two experiments: the first used a bilingual Stroop task in which English and French colour terms appeared in congruent (i.e. the word “red” presented in red font) and incongruent (i.e. the word “red” presented in green font) conditions. This allowed us to examine linguistic processing. The second experiment used the non-linguistic ANT task in which sets of 5 arrows appeared with the central target arrow in either a congruent (e.g. ← ← ← ← ←) or incongruent (e.g. ← ← ← ← ←) condition in order to examine non-linguistic processing. Reaction time (RT) data were recorded and analysed.

In an analysis of reaction times, a one-way ANOVA revealed an unexpected lack of differences in the magnitude of Stroop interference between the Functional Monolingual, Simultaneous, and Early bilingual groups. For the ANT task, however, a significant effect for the attentional network of orienting was found (\(p = .011\)). A planned pair-wise comparison revealed that Early bilinguals had significantly better attention orienting abilities than did Simultaneous bilinguals (\(p = .010\)), but there were no significant differences between the Functional Monolinguals and Simultaneous bilinguals.

These findings demonstrate non-linguistic cognitive control differences between the two groups of immersed bilinguals, which may reflect how they acquired and thus process their L2. Because the Simultaneous bilinguals acquired both languages at the same time, their two languages may have been integrated from the start. As such, there is no conscious processing conflict to mediate, which is said to be the basis for the bilingual advantage. For Early sequential bilinguals, however, the L2 has to be integrated with an already-existing L1; thus, the Early bilingual must put to use cognitive control skills to mediate the conflict between the L1 and L2. Therefore, Early bilinguals may enjoy greater cognitive control advantages in non-linguistic tasks, even if their language processing abilities are similar to those of Simultaneous bilinguals who became immersed in their “L2” at an earlier age.


EFFECTS OF LANGUAGE INPUT REGULARITIES THROUGHOUT DEVELOPMENT
Marius Janciauskas (University of Liverpool), Franklin Chang (University of Liverpool)
m.janciauskas@liv.ac.uk

Language learning and adaptation theories often assume that linguistic input has similar effects on different populations (e.g., adults, children, L2 learners). However in L2 learners, in addition to an overall drop in performance as people’s age of language acquisition (AoA) increases, the performance with different grammar rules drops at different rates (Johnson & Newport, 1989). Also, some studies have failed to find the effect of input measures like length of L2 exposure (LoE) on grammar knowledge at all (J&N’89). These phenomena need to be explained in order to understand how mechanisms of language learning/adaptation interact with input throughout development.

We hypothesized that variation in performance with different rules at different AoA arises from frequency effects in learners’ input. To test this, we computed conditional probabilities (CP; e.g. frequency of past tense verbs relative to overall verb frequency) for 10 grammar rules based on the test items in J&N’89 study in the CHILDES corpora (MacWhinney, 2002). We then reanalysed Flege et al.’s (1999) sentence grammaticality judgment scores of 240 Korean learners of English. We used AoA crossed with LoE (years) and rule CP to predict correct responses in a logistic mixed model. We found a negative effect of AoA confirming an overall ‘sensitive period’ effect. We also found a positive effect of LoE but it was evident only in early (<12 years) AoA learners (Fig. 1; shows 2 rules for simplicity). This can explain why previous studies have failed to find strong correlation between LoE and L2 grammar knowledge. We also found that early learners improved more with higher CP rules (plural (PLU) vs. past tense (PST)) but this was not evident in late learners. Finally, late AoA learners showed worse performance with higher CP rules, which suggests that greater variation in frequent rules can make it harder to reject incorrect test items. These findings suggest that language-learning mechanisms undergo changes that compromise late learners’ abilities to use distributional information in the input.

To better understand these findings, we modelled the results using a connectionist model of language acquisition and production (Chang, 2009). It was first trained using a Korean language for various lengths of time before introducing English as a second language. This resulted in 16 models with different AoA. We found that the models with a constant learning rate exhibited sensitive period effect on the overall model accuracy due to entrenchment from L1 language. However, gradual reduction of learning rates was necessary to model the effects associated with rule CP like greater effect of LoE on higher CP items or the negative effect of higher CP rules on performance in later AoA models. These findings support the idea that input-driven language learning mechanisms undergo changes that affect the way these mechanisms interact with input regularities at different developmental stages. Further work is needed to identify the factors that influence these changes.

References:
How young children come to represent syntax abstractly is fundamental to how humans evolved language and how it is learned. Evidence on this varies. Children as young as 3-4 years old demonstrate priming of abstract syntactic structure about as much as adults (Rowland et al., 2012), suggesting that children represent syntax abstractly from an early age. On the other hand, 2-3 year old children do not readily generalize between structural alternates for novel verbs (Brooks & Tomasello, 1999), suggesting that their syntactic representations are to some extent tied to the lexical content of the sentences in which they occur.

One explanation for these disparate findings is that young language learners represent syntax abstractly but are hesitant to generalize this knowledge to novel stimuli because certain verbs (e.g. “donate”) and certain events (e.g. transfer to a non-sentient recipient) license only one member of a syntactic alternation. Thus, children may tend to maintain the verb-syntax pairings used by other speakers until they are confident they know which restrictions apply to a particular verb or event.

In the present study, we explored the tendency of adults and 4-5 year old children to repeat a confederate’s syntactic structure to describe a particular event (syntactic entrainment). First, confederates described a set of 6 action pictures. Each picture was described with an active or a passive, a prepositional or a double object dative, or a with- or an in-locative. Next, subjects described the same pictures back. We measured whether subjects used the same structure to describe a picture as they heard the confederate use to describe the picture initially. Additionally, target pictures were accompanied by specific verbs, which were the same or different as the one used by the confederate. If children represent syntax abstractly they should repeat the confederate’s syntactic structure whether or not the verb is repeated. But if they hesitate to generalize, they should repeat the syntactic structure even more when they also repeat the verb, where they can rely on the confederate’s verb-syntax pairing.

Overall, both children and adults were more likely to use the same syntactic structure as the confederate (21% and 16% respectively), as compared to the alternative structure. Adults’ repetition of syntax was not affected by whether they used the same verb (15%) or a different verb (16%) as the confederate. However, children were more likely to repeat the confederate’s syntax when they used the same verb (29%) than a different verb (13%). The difference of the effect in the same verb condition compared to the different verb condition was marginally significantly greater for children than adults ($F = 3.27; p = .075$).

Both children and adults seem to be able to represent syntax abstractly—repeating the confederate’s syntax even when using a different verb. Unlike adults, however, children show a greater tendency to repeat syntactic structures when the verb is repeated. This indicates that they assume a more “concrete” view of verb-syntax pairings when used to describe particular novel events. This potentially reconciles the divergent findings on abstract syntax in toddlers. They have abstract syntactic representations, but they are still figuring out when and how to use them.


MULTIPLE SYNTACTIC PROCESSING BIASES THAT CHANGE OVER DEVELOPMENT
Kirsten Abbot-Smith (University of Kent), Franklin Chang (University of Liverpool), Caroline Rowland (University of Liverpool), Heather Ferguson (University of Kent) & Julian Pine (University of Liverpool)
K.Abbot-Smith@kent.ac.uk

Many preferential-looking studies have found that young children will map an active transitive sentence containing a novel verb onto a causative event in which the first noun is the agent (e.g., Gertner et al., 2006). Gertner and Fisher (2012) argue that one and two-year-olds will wait until the end of the sentence before assigning semantic roles in this manner. In contrast, theories that link adult sentence processing and acquisition propose the children process even sentences with novel verbs incrementally and English-speaking preschool children initially have both an incremental bias to map the first noun onto agent during the first noun phrase as well as post-verb processing biases (Chang, Dell, & Bock, 2006). To compare these two accounts, we carried out an eye-tracking study of comprehension of active transitive and full passive sentences containing novel verbs by 2-year-olds, 3-year-olds and adults. We adapted the paradigm used by Gertner et al. (2006: Study 3) whereby participants simultaneously see two video clips of novel causative events, both involving a boy and a girl, whereby in one clip the boy is the agent and in the other he is the patient. For our main analyses (Figure 1), we examined 800 msec windows that were timelocked to both first and second NP taking into account the mean lexical processing speed for each age group. In the first noun region, the children showed a bias to map the first noun onto the agent, both for active and passive sentences, but this bias was reduced in adults. In addition, the second NP region showed a differentiation between active and passive structures, and this differentiation grew stronger over development.

Figure 1: Mean proportion looks to the active-target clip by age, region, and structure

Thus, the first noun as agent bias is an incremental one and decreases with age. Three-year-olds clearly show multiple position-specific bias and are able to recover from an initial garden-path (in the form of a first noun as agent bias) in their processing of full passive sentences with novel verbs. These results support theories that argue for continuity between children and adults in that sentence parsing is incremental and makes uses of multiple learned biases between meaning and form.
THE ROLE OF DIFFERENT TYPES OF CUES IN DISAMBIGUATING NON-CANONICAL SENTENCES

Valentina Cristante (University of Münster and Osnabrück)
vcristante@uni-osnabrueck.de

In the literature on child sentence processing, it is unclear when children start attending to a cue and how much weight they will attribute to it (e.g. Dittmar et al., 2008; Choi and Trueswel, 2010). The goal of this study is to observe how the position, the kind of the disambiguating cue and the type of reanalysis required for a correct understanding of the sentence influence the comprehension and the processing of non-canonical sentences. To this aim, the standard canonical surface structure Subject-Verb-Object (SVO) in German is compared to three different non canonical structures presenting disambiguating cues which have different positions in the sentence and a different form. This study investigates how monolingual German children (n=44) and Turkish-German children (n=37) (mean age=7,1, age of L2 onset= 3 to 4 years old) perform during a sentence-picture matching task while hearing canonical unambiguous SVO sentences (1), SVO passive sentences (2) disambiguated by the auxiliary and the prepositional “by phrase”, Object-Verb-Subject (OVS) sentences presenting already a clear disambiguating case marking on the 1st (male) NP (3) and OVS sentences with female case ambiguous initial NPs and disambiguating male 2nd NPs (4). While participants listened to the sentences, their eye-movements to the two possible pictures were monitored and their comprehension accuracy was measured at the end of the sentence.

(1) SVO, first NP male -Der Opa hat am Mittag die Oma (Obj.F.amb.) aus Versehen geweckt
(2) SVO, Passive, first NP male -Der Opa (Subj.M.not amb.) wurde am Mittag von der Oma (…) 
(3) OVS, first NP male -Den Opa (Obj.M.not amb.) hat am Mittag die Oma (Subj.F.amb.) (…) 
(4) OVS, first NP female -Die Oma (Obj.F.amb.) hat am Mittag der Opa (Subj.M.not amb.) (…) 

<table>
<thead>
<tr>
<th></th>
<th>SVO, NP1 male</th>
<th>SVO, Passive, NP1 male</th>
<th>OVS, NP1 male</th>
<th>OVS, NP1 female</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Children</td>
<td>92,5</td>
<td>86,4</td>
<td>55</td>
<td>8,75</td>
</tr>
<tr>
<td>L2 Children</td>
<td>80,6</td>
<td>70,8</td>
<td>29,5</td>
<td>12,5</td>
</tr>
</tbody>
</table>

The accuracy results show that passive sentences do not represent a strong problem for both L1 and L2 children. In contrast, OVS sentences do, in particular during OVS, NP1 female, where the disambiguating marker appears later in the sentence. During OVS sentences with a male first NP, L1 children perform slightly above chance and do not show adult like patterns, contrary to what was observed in previous studies (e.g. Dittmar et al., 2008). The low score of L2 children in this condition (29,5%) reveals their stronger reliance on word order than on case marking. The eyetracking data allow us to observe how the participants parse the different morphosyntactic cues without the interference of other factors involved in the offline task (Marinis, 2010). In the passive condition, both children groups adopt a first NP strategy that they abandon after having encountered the auxiliary “wurde” (L2 children are slightly slower in the reanalysis). This does not happen for OVS sentences, where L2 children in both conditions tend to strongly over-rely on word order with a minimal reaction to the disambiguating cues. In contrast, the L1 children are more sensitive to case marking, reacting to the disambiguating accusative “den” on the first NP, but still revealing strong difficulties in revising OVS sentences with female first NP where the disambiguating cue appears late in the sentence. The results show how the position, the type of cue (auxiliary or case marking) and the kind of reanalysis triggered by a cue (reassignment of semantic roles in the case of passive, and of both semantic and syntactic roles in the case of OVS sentences) influence the comprehension and the processing of non-canonical sentences.
THE PROCESSING OF MORPHOSYNTACTIC CUES IN WH-QUESTIONS: 
AN EYE-TRACKING STUDY 
Atty Schouwenaars (University of Oldenburg), Esther Ruigendijk (University of Oldenburg) &  
Petra Hendriks (University of Groningen)  
atty.schouwenaars@uni-oldenburg.de

Many studies report that object questions, in which the object precedes the subject, are  
difficult to comprehend for children due to a strong subject-first bias (among others; [1]). The  
ability to identify who is doing what to whom in sentences with a non-canonical word order by  
the use of case-marking does not develop until the age of 7 ([2]). Children seem to be even  
less sensitive to the verb agreement cue ([3] for Dutch, [4] for Italian). 

The aim of this study is to find out (1) when and to what extent German children make use 
of these cues in their interpretation of object questions, and (2) whether different cues lead to 
different processing (also for adults), in terms of different eye-gaze patterns.

We tested 7-10 year old German children (n=33), and an adult reference group. A picture 
selection task with eye-tracking was carried out to test the comprehension of wh-questions  
(subject, object and passive questions) disambiguated by different types of morphosyntactic 
cues; agreement and case (AgCa), agreement only (Agre), and case only (Case), see Fig. 2.

Offline data confirm that children (as well as adults) respond less correctly on object  
questions (86% correct), than on subject (98%) and passive questions (98%). The 7 year old  
children respond less accurately (70%) on object questions than older children. Offline data  
do not indicate differences between the different types of cues, but online data do: The eye- 
gaze patterns show that for both adults and children, object questions disambiguated later in  
the sentence by agreement and case cues (on the article of the second NP) lead to incorrect  
first interpretations, due to a subject-first bias (see Fig.3). Object questions disambiguated by 

Case-marking on the wh-element do not lead to incorrect first interpretations.

Our eye-tracking study shows that German children are sensitive to both case marking  
and verb agreement and do revise their incorrect first interpretations of object questions.

Fig. 1: Example of picture pair with reversible roles  
Fig. 2: Examples of object questions with different cues

Fig. 3: Gaze data Object questions Children

The comprehension of wh-questions in three subtypes of SLI. Lingua, 121(3), 367-382. [2]  
comprehension of word order and case marking in causative sentences. Child Development,  
which questions in five to nine-year-old Dutch children: The role of number. New developments in  
comprehension of interrogative sentences. In: Bagnara, S. (Ed.), Proceedings of the European  
The Acquisition of L1-L2 Grammatical Gender in German-French and English-French Bilinguals

Leah Pappas (Pennsylvania State University), Carrie N. Jackson (Pennsylvania State University), Janet van Hell (Pennsylvania State University; Radboud University Nijmegen) & Katharina Spalek (Humboldt-Universität zu Berlin)

A growing number of studies have reported that first (L1) and second (L2) language gender systems interact during L2 production (e.g., Bordag & Pechmann, 2007). Studies show that when genders are the same across languages, the L1 can facilitate L2 gender processing, and it can inhibit processing when genders conflict (e.g., Lemhöfer et al., 2008; Paolieri et al., 2010). However, questions still exist as to whether these influences stem from difficulties in acquiring accurate L2 gender representations or the inhibition of L1 gender during online L2 processing (e.g., Lemhöfer et al., 2008). The present study uses reaction times (RTs) to investigate L1 influence on L2 gender processing when there is no gender equivalent between the languages – an area that has not yet been extensively explored. The focus is the neuter gender in German, which has no French equivalent, and therefore, may not provide direct competition during L2 gender production in German-French bilinguals. Testing English-French bilinguals of similar L2 French proficiency provides a baseline for comparison, since English does not have grammatical gender.

Fifteen German and 14 English advanced L2 speakers of French completed two L2 picture-naming tasks: a bare noun naming task followed by a determiner-noun naming task. Stimuli varied according to congruence (same gender in L1 and L2), incongruence (opposite gender between the L1 and L2), and no gender match (neuter in German).

<table>
<thead>
<tr>
<th></th>
<th>Congruent</th>
<th>Incongruent</th>
<th>No Gender Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>la.f fleur</td>
<td>le.m pont</td>
<td>la.f feuille</td>
</tr>
<tr>
<td>German</td>
<td>die.f Blume</td>
<td>die.f Brücke</td>
<td>das.n Blatt</td>
</tr>
<tr>
<td>English</td>
<td>the flower</td>
<td>the bridge</td>
<td>the leaf</td>
</tr>
</tbody>
</table>

Statistical analyses using mixed-effect models reveal that RTs for all three conditions in the bare noun task were longer than RTs in the determiner-noun task, with the longest RTs occurring on incongruent items for both groups, suggesting that the bare noun task served as a familiarization phase for the participants. More critically, in the determiner-noun task, the German-French bilinguals named items with congruent gender across their two languages faster than incongruent items. No Gender Match items elicited RTs that were between those of the congruent and incongruent categories; they also showed the highest naming accuracy. In contrast, the English-French bilinguals exhibited no significant RT or naming accuracy differences across conditions in the determiner-noun naming task. Among the German-French bilinguals, the lack of significant RT differences between the Congruent and No Gender Match items, as compared to the significantly longer RTs on Incongruent items, suggests that for L2 speakers with grammatical gender in their L1, competition between L1 and L2 gender may be more limited when there is no direct gender equivalent in the L1 and the L2.

REFERENCES


Do the effects of bilingualism in late unbalanced bilinguals and multilinguals mirror those of early bilinguals? Previous studies have reported both costs and benefits of early bilingualism on different cognitive domains: on one hand, bilinguals are reported to be slower on lexical access tasks; on the other hand, they seem to perform better on tasks involving executive functioning [1]. We explored whether bilingualism would impact cognitive performance of late-unbalanced bilinguals and multilinguals in a similar fashion as early bilingualism. To address this issue, we conducted two experiments using a task that has been shown to have cognitive costs for bilinguals (picture-name matching task) and a series of non-linguistic tests of executive control in which bilinguals have shown a better performance (sustained attention, selective attention and attentional switching). One-hundred and ninety three native speakers of English took part in these studies. Participants were divided into three groups: monolinguals, bilinguals (with knowledge of one foreign language) and multilinguals (with two or more foreign languages). The bilinguals and multilinguals had Spanish or Italian as their dominant foreign language and we assessed the proficiency in that language through the picture-name verification task, as well as a self-rating scale.

Results from Experiment 1 showed that bilinguals, and particularly multilinguals, were slower to respond to L1 words than monolinguals on the picture matching task (Fig. 1). In contrast, both bilinguals and multilinguals outperformed monolinguals on the selective attention task. We then asked how much exposure to the L2 would be necessary for these effects to appear. Using the same methodology as in Experiment 1, we conducted Experiment 2, controlling for different levels of proficiency. We divided the three groups into years of study into first and fourth (final) year of university. Additionally, we collected information about the years of musical experience which has been suggested to have an impact on performance on executive control tasks. Results of Experiment 2 replicated our previous findings: bilinguals and multilinguals showed slower reaction times to native words, relative to monolinguals. On the executive control tasks, bilinguals and multilinguals outperformed monolinguals already in their first year of study on the test of selective attention. This difference remained significant in the fourth year, despite an overall improvement in performance across all three groups. On the attentional switching task, a bilingual and multilingual advantage was detected only in the fourth year (Fig. 2). All these effects were independent from the years of musical experience. Our results suggest that late-acquisition non-balanced bilinguals and multilinguals experience similar cognitive effects as their early-acquisition, balanced counterparts and that some of these effects appear already at an early stage of second language acquisition.

How do different forms of semantic knowledge impact incremental interpretation in native and non-native speakers? Some researchers have argued that semantic memory organization (in particular, semantic relatedness) overrules early effects of sentence plausibility as long as coarse-grained semantic expectations about animacy are met ([1]; henceforth P&K). Here, we investigated effects of semantic relatedness on how people process plausibility violations and animacy violations, using materials based on P&K. We also examined whether the interaction between animacy and plausibility constraints is different in bilinguals whose native languagemarks animacy grammatically. Twenty L1-English speakers and twenty Spanish L2-English bilinguals read passive sentences word by word (450ms duration, 100ms blank, post-sentence yes/no plausibility judgment) with sentential subject NPs being (C) plausible-animate, or implausible and (RA) related-animate, (UA) unrelated-animate, (RI) related-inanimate or (UI) unrelated-inanimate (Fig. 1a). Items were pre-tested on plausibility by both groups to match P&K (31 items per condition, 90 fillers to balance plausibility), LSA-SSVs indexed relatedness. ERP analyses focused on 350-450 (N400) and 700-900 (P600) ms time windows, using 2x2 repeated measures ANOVAs with factors animacy and semantic relatedness tested at 16 electrode sites (divided in 4 ROI-quadrants by Anterior/Posterior and Left/Right). Only significant findings are reported.

Monolinguals showed larger N400s for inanimate NPs than for animate NPs and for unrelated NPs compared to related NPs (Fig. 1b), thus resulting in two main effects rather than in the interaction observed in P&K. The P600 window revealed a two-way relatedness by AP distribution interaction: related NPs elicited larger P600s at anterior regions than unrelated NPs. Importantly, monolinguals showed larger animacy effects than bilinguals. The bilinguals showed larger N400s for inanimate NPs but not for unrelated NPs, while showing the same interaction as monolinguals in the P600 window.

Our results partially replicate P&K’s, as we also found largest N400 effects for animacy violations. However, monolinguals showed an additive effect of semantic relatedness, rather than the interactive pattern reported by P&K with only a relatedness effect for animate NPs. The smaller N400s for related NPs irrespective of animacy suggests that semantic relatedness facilitated processing incoming NPs regardless of the fulfillment (or lack thereof) of animacy expectations. The L1 N400 results may reflect coarse-grained expectations involving animacy and relatedness, or, alternatively, online effects of subtle differences in plausibility. In contrast, the L2 N400 results could reflect that bilingual predictions are primarily driven by animacy, or reduced sensitivity to subtle differences in plausibility despite sensitivity to general implausibility.

Recent years have seen an expanding size and scope of studies on the plasticity of the brain associated with learning involving higher-level brain functions such as language acquisition. Not much is known, however, regarding the neural basis of changes in learning mechanisms and how these might be modified by maturation and experience, despite the fact that the period from late childhood to young adult has long been viewed as a period accompanied by significant mental development and change in learning mechanisms.

In the present study, we examined functional changes and neuroplasticity of 71 participants aged 9 to 21 years old using behavioral and brain-function measurements involving near-infrared spectroscopy, while half the participants engaged in approximately one-month-long training for acquiring phonemic distinctions in a non-native language (English) and the other half engaged in a control task.

Figure 1 shows the behavioral result which indicated that every age group that engaged in phonemic learning did achieve significant improvement ($p < .0001$) in their ability to distinguish English phonemes unlike the control group ($p = .87$), as indicated by the higher slope value in the phonemic identification curve.

Using the same method that is used in previous studies (e.g. Fuchino et al., 2013⁵), we examined temporal correlations of activity between left Supramarginal Gyrus (lSMG) (which is said to be crucially involved in phonological processing) and temporal and frontal cortical regions during categorization tasks using near-infrared spectroscopy. In the Figure 2, circles whose circumferences are black correspond to those regions whose z values exceed zero. In agreement with prior studies, functional network in children’s brains is more widespread but not so much in frontal regions, while adults (and adolescents) rely more on frontal regions, probably because the bottom-up learning of the sensitive period becomes increasingly influenced and gated by top-down processes in later years.

These results indicate that learning directed toward acquisition of spoken language causes changes in behavior and brain function in all age groups and that children on the one hand and adolescents and young adults on the other show distinctive patterns of changes in functional connectivities.

---

Reference: doi:10.1371/journal.pone.0067432
Previous research has shown that from 7.5 months of age infants use rhythmic patterns to extract words from speech, relying on stressed syllables in English or German (Jusczyk et al., 1999, Bartels et al., 2009). At the same time, artificial language studies show that 7.5-month-old infants exploit pitch in order to group disyllables (Bion et al., 2011), but already at 9 months, they generalize over pitch in word recognition tasks, at least in English (Singh et al., 2008). In two head-turn preference experiments, we tested infants' reliance on pitch and metrical stress in segmentation.

In Experiment 1, German 10-month-olds were familiarized with two trisyllabic target words with penultimate stress (e.g., Lagune), embedded 6 times in short text passages. Target words were produced with a pitch fall, whose peak was either realized before (H+L*, 'early peak') or on the stressed syllable (H*, 'medial peak') (Kohler, 2012). The naturally recorded target words were matched for pitch range, duration, and vowel quality across intonation conditions. For the test lists, the speaker recorded 15 repetitions of trochaic part-words starting at the stressed syllable (e.g., gune) with a high-falling pitch accent. Infants were randomly assigned to 2 of the 4 familiarization paragraphs; the familiarization phase ended when the infant had attended to each of the 2 paragraphs for at least 45s. During test, infants listened to 3 repetitions of the 4 test lists, each containing 15 tokens separated by 800ms of silence in a randomized order. The duration of orientation for each test list was recorded on-line and was averaged for novel and familiar lists for each infant.

Log-normalized looking times for 32 infants showed a significant interaction between familiarity and intonation condition (p<0.05). There was a novelty effect in the medial peak condition only (p<0.05): infants looked 1.3s longer to novel than to familiarized part-words.

In Experiment 2, we tested whether infants recognize the part-words in a different intonation contour. We familiarized another group of 17 infants with the paragraphs in the medial peak condition and tested them on the same trochaic part-words, but this time produced with a pitch rise instead of a fall. Results also showed a novelty effect of 1.2s (p=0.05). Our results demonstrate that German 10-month-olds rely on high-pitched syllables for lexical segmentation. They fail to segment when the pitch peak precedes the stressed syllable. They also generalize over the intonational realization in recognition, however, suggesting that intonation only plays a role for segmentation, but not in early lexical representations.


INTERACTIVE PARTICIPATION FACILITATES CHILDREN’S PROCESSING OF DIRECT SPEECH REPORTS

Franziska Köder (University of Groningen)
f.koder@rug.nl

Direct and indirect speech differ with respect to the perspective from which a previous utterance is reported. In indirect speech (example (1)), the reporting speaker is consistently the deictic centre. However, in direct speech (example (2)) the perspective shifts from the reporting speaker to the reported speaker, inducing a shift in the interpretation of pronouns like I.

(1) Indirect speech: Elephant said that I get the car. (Dutch: Olifant zei dat ik de auto krijg.)
(2) Direct speech: Elephant said, “I get the car.” (Dutch: Olifant zei, “Ik krijg de auto”.)

There is an apparent paradox in the development of reported speech: From a semantic perspective, we expect direct speech to be harder to process because of the perspective shift (Kaplan, 1989). On the other hand, direct speech is cross-linguistically more widespread (Evans, 2013), acquired earlier and used more frequently by children (Ely & McCabe, 1993; Nordqvist, 2001). In experiment 1, we provide support for the semantic-based explanation showing that pronoun processing is harder in direct speech. In experiment 2, we show that children’s performance on direct speech increases dramatically if the child experiences the situation not as external observer but as interactive participant.

Experiment 1: 136 children aged 4, 5, 6, 7, 9 and 11 and 33 adults – all native speakers of Dutch – took part in experiment 1. The experiment is designed as a tablet game called ‘which animal gets which object’. In each scene, one animal whispers – inaudibly for the participant – into another animal’s ear which of the three animals gets a certain object. The original addressee tells the information then to the third animal using either Dutch direct or indirect speech with a first (I), second (you) or third (he) person singular pronoun (see examples (1) and (2)). The participant’s task is to select the recipient of the object, i.e. the referent of the pronoun. We found that in all age groups participants make significantly more mistakes and have longer reaction times for the interpretation of pronouns in direct speech as compared to indirect speech. Age has a positive effect on accuracy, but even 9- and 11-year-old children struggle with the perspective shift in direct speech (see Fig. 1 for the 9-year-olds (N=20)).

![Figure 1: Experiment 1 (external observation): 9-year-olds](image1)

![Figure 2: Experiment 2 (interactive participation)](image2)

Experiment 2: In experiment 2, we investigated whether involvement in the interaction increases children’s comprehension of direct speech. The procedure is similar to experiment 1. However, this time the participant did not watch the interaction as an external observer on a tablet but directly interacted with two gender-matched hand puppets. After hearing a direct or indirect speech report, the participant had to put objects in the correct box (her own or that of hand puppet 1 or 2). 22 8-year-olds took part in experiment 2 (see results in Fig. 2). In comparison with the 9-year-old children from experiment 1 (Fig. 1), their interpretation of pronouns in direct speech is significantly better. However, not all pronouns are equally hard to interpret. The second person pronoun is the most difficult one in direct speech because children have to inhibit the tendency to interpret it as referring to themselves. Children’s improvement in experiment 2 could be due to two factors. First, children are involved in the interaction instead of being external observers (Charney, 1980; Murphy, 1986). Second, speech reports are situated in a more naturalistic environment involving perception and action (Barsalou, 1999).
ORTHOGRAPHIC PROCESSING IN BILINGUAL CHILDREN: 
THE ROLE OF SEMANTICS IN COGNATE RECOGNITION 
Pauline Schröter & Sascha Schroeder (Max Planck Institute for Human Development, Berlin) 
schroeter@mpib-berlin.mpg.de 

Lexical access in bilinguals is assumed to be language-nonselective and based on an integrated lexicon. The BIA+ (Bilingual Interactive Activation) model (Dijkstra & Van Heuven, 2002) postulates that the visual presentation of a word leads to co-activation of orthographic representations in the two languages of a bilingual. In succession, associated phonological and semantic representations are activated, the language is selected and a lexical candidate is identified. A large body of research on bilingual processing has shown that words which share the same spelling and meaning across languages, so-called cognates, are recognized faster than matched control words, which has been attributed to stronger activation caused by their shared orthographic entry in the bilingual lexicon. However, studies on words that are identical in spelling but different in meaning, so-called false friends, have revealed that shared orthography alone does not lead to facilitation (e.g. Lemhöfer & Dijkstra, 2004). This, in turn, argues for semantic-to-orthographic feedback in bilingual word processing and could be interpreted in line with the assumption of the BIA+ model that, unlike cognates, false friends are represented by different entries in the bilingual lexicon. Yet, little is known about cognate processing in children and the development of the bilingual lexicon. Studies on child second language learners have found cognate facilitation effects only in their second language (Brenders, van Hell, & Dijkstra, 2011), which is argued to be due to the imbalance in language proficiency. In balanced bilingual children, therefore, cognates should lead to facilitation irrespective of the language. Also, cognates should only differ in their recognition from false friends if semantic co-activation is already present at the beginning of bilingual reading acquisition. However, it has never been investigated whether predictions made by the BIA+ model for bilingual adults also generalize to children.

We investigated the processing of cognates and false friends in balanced German-English bilingual children. Students aged 8-9 years, who had learnt to read simultaneously in both languages, performed a lexical decision task in German and English. Cognates and false friends were matched on word length and frequency and counter-balanced in their presentation across languages. Control words were translation equivalents and likewise matched on length and frequency. Results showed a strong processing advantage for cognates over controls in both languages, indicating language-nonselective lexical access already in the early stages of reading acquisition. False friends did not differ in their recognition from false friends if semantic co-activation is already present at the beginning of bilingual reading acquisition. In sum, results indicate that with respect to orthographic processing balanced bilingual children are more comparable to bilingual adults than to child second language learners. Findings are interpreted in terms of theories on reading acquisition and developmental aspects within the BIA+ framework.

References:
Inflectional morphology has proven a test case for evaluating different theoretical approaches to language processing. Most of the debates have been based on data acquired using the inflection from stem task, but it is not universally accepted that this accurately approximates more naturalistic speech production. Indeed, work with adults reveals a strong effect of task context on skilled inflection, such that the robust regularity effect seen after presentation of a stem is eliminated when the stimulus is instead an action picture (Woollams et al., 2009). A large scale connectionist computational model trained on multiple tasks (repetition, hearing, speaking and inflection) was able to accurately simulate these task differences in the presence of the regularity effect.

The standard “wug” task (Berko, 1958) used to study morphological development also involves stem presentation, yet the impact of task context upon the trajectory of past-tense inflection has not been explored. We recently assessed task differences in 900 children aged between 2;6-5;5. Children completed either a standard wug-task in which they heard a verb modelled in its progressive form and were prompted to produce a past form, or a video-task in which no prior verb model was provided. The same 162 verbs were elicited in both tasks, with each child receiving a subset. As shown in the left graph below, children of all ages showed a clear reduction in the size of the regularity effect in the video-task relative to the wug-task, although the regularity effect remained significant for all cases in both tasks.

We explored the ability of a version of the connectionist model of inflection to simulate these developmental data. The model consisted of distributed representations of input and output phonology linked via a hidden layer that connects to distributed semantic representations. The model was trained using frequency weighted presentation of 785 monosyllabic verbs that occur in the child directed speech of the Manchester Corpus (Theakston et al., 2001). The model began with comprehension trials, then speaking trials were introduced on the fifth training epoch and inflection trials on the fifteenth. Weight decay in the model was gradually decreased over training to emulate more efficient learning with age. The wug-task was simulated by activation of the representation of the progressive form over the input phonological units, while the video task was simulated by activation of the representation of the meaning of the verb over the semantic units. Model performance (averaged over 25 instantiations) was considered from the point in training comparable to the youngest children in terms of accuracy for regular verbs in the wug-task. As can be seen in the right graph below, the model’s accuracy on the monosyllabic subset of the verbs used with the children reflected the same pattern of reduced regularity effects with when inflection was in response verb meaning rather than the progressive phonological form.

The behavioural results show a clear influence of task context on children’s performance and our simulations demonstrate the potential for such differences to be captured by a multiple task connectionist model in which learning occurs through exposure to a representative corpus of verbs.
PERSPECTIVE-TAKING IN TEXT IS MODULATED BY HOW INFORMATION IS PRESENTED (DIRECT VS. INDIRECT SPEECH)

Steven Frisson (University of Birmingham), Elisa Back (Kingston University), Alice Rees (University of Birmingham), & Ian Apperly (University of Birmingham)
s.frisson@bham.ac.uk

Recent evidence (Yao & Scheepers, 2011; see also Stites, Luke, & Christianson, 2013) has shown that readers adapt their reading rate of direct or quoted speech information according to how fast the speaker is supposed to have uttered the quote. However, when the same information is presented as indirect or non-quoted speech (using that-clauses), this effect disappears. What has not been examined though is whether information in direct and indirect speech is processed in the same way and to the same extent.

We present two eye movement experiments investigating perspective mismatch effects during normal text reading. In the experiments, readers are presented with a narrative introducing two characters and two objects (e.g. apple and pear) or actions (e.g. playing football or playing tennis). While one character was aware of both objects/actions, the other was only aware of one. In the final target sentence, one character would comment on one of the objects/actions. Crucially, the sensicality of the target sentence depended on the character and the object/action, and readers are expected to notice a mismatch if they calculated the character’s perspective on-line. Example:

Billy and his mum go to a theme park. Billy wants to go on the roller coaster but it is too scary for his mum so she lets Billy go on it by himself and arranges to meet him afterwards at the ice cream van. After the roller coaster Billy sneaks a ride on the ghost train. He then goes to meet his mum.

Target:

a/b His mum says “You were very brave to go on the ghost train/roller coaster as it…”
c/d His mum thought he was very brave to go on the ghost train/roller coaster as it…

We found evidence of early disruption (first-pass duration) of perspective mismatch (a&c longer than b&d), indicating that the reader had kept track of the character’s knowledge base. In addition, later measures showed a stronger mismatch effect for the non-quoted/indirect speech condition than the direct speech condition.

Experiment 2 used additional stimuli and a larger number of participants. The results showed that the mismatch effect was only present in the indirect speech condition but not in the direct speech condition. Together, the data suggest that readers are more likely to notice a perspective mismatch when the information is presented as non-quoted speech. We interpret this finding with respect to depth-of-processing (e.g. Good Enough theory, Ferreira & Patson, 2007) and the relative difficulty of making first-order (for indirect speech) and second-order (for direct speech) inferences.


IRONY AND EMOTICONS: COMPREHENSION AND EMOTIONAL IMPACT
Ruth Filik, Alexandra Turcan, Dominic Thompson, Nicole Harvey, Harriet Davies (University of Nottingham) & Hartmut Leuthold (University of Tuebingen)
ruth.filik@nottingham.ac.uk

The conventional view of verbal irony is that it is a figure of speech in which the speaker says the opposite of what they actually mean. For example, if someone does something stupid, their friend may make the ironic comment “That was clever!”. Given the risk of misunderstanding, it is not clear why people choose to use ironic expressions instead of speaking literally. However, most theorists would agree that it may serve some other communicative function that would not be achieved by speaking directly, such as eliciting a particular emotional response in the recipient. One point of controversy concerns whether ironic language is used to enhance (e.g., Colston, 2007) or to mute (e.g., the Tinge Hypothesis, Dews & Winner, 1995) the positive or negative nature of a message, compared to literal language. Previous research highlights the potential role of emoticons and other textual devices in the interpretation and emotional impact of ironic utterances (e.g., Derks, Bos, & Von Grumbkow, 2008; Hancock, 2004; Wolf, 2000), however, it is currently unclear which devices are most effective. Thus, the aims of the current research are to investigate the comprehension of, and emotional responses to, ironic vs. literal criticism and praise, and the influence of textual devices on both comprehension and emotional impact.

144 participants read 48 materials, and rated them in terms of how ironic the final comment appeared to be (“not at all ironic” to “very ironic”), and how they thought the recipient of such a comment would feel (“very negative” to “very positive”). Textual devices investigated were: the tongue face emoticon :p, the wink face emoticon ;-) and ellipsis (...).

Ironic criticism (with device): Tanya had noticed that Jenny had put on a lot of weight. She texted her to say, “I see the diet is going well :-p / ;-) / …”
Ironic criticism (no device): Tanya had noticed that Jenny had put on a lot of weight. She texted her to say, “I see the diet is going well”

Literal criticism (with device): Tanya had noticed that Jenny had put on a lot of weight. She texted her to say, “I see the diet is going badly :-p / ;-) / …”
Literal criticism (no device): Tanya had noticed that Jenny had put on a lot of weight. She texted her to say, “I see the diet is going badly”

Ironic praise (with device): Tanya had noticed that Jenny had lost a lot of weight. She texted her to say, “I see the diet is going badly :-p / ;-) / …”
Ironic praise (no device): Tanya had noticed that Jenny had lost a lot of weight. She texted her to say, “I see the diet is going badly”

Literal praise (with device): Tanya had noticed that Jenny had lost a lot of weight. She texted her to say, “I see the diet is going well :-p / ;-) / …”
Literal praise (no device): Tanya had noticed that Jenny had lost a lot of weight. She texted her to say, “I see the diet is going well”

In terms of comprehension, for literal comments, both criticism and praise were judged as more ironic when a device was present (for all three devices). For ironic comments, when the tongue face or wink face emoticon were absent, criticism was rated as more ironic than praise but both were rated as equally ironic when the emoticon was present. In contrast, for ellipsis, there was no difference in irony ratings between criticism and praise with or without the device. Regarding emotional impact, results overall supported the Tinge Hypothesis: ironic criticism was rated as less negative than literal criticism and ironic praise less positive than literal praise. In addition, both the tongue face and the wink face emoticon made ironic items (criticism and praise) appear more positive, and also made literal criticism appear more positive (but had no effect on literal praise). The ellipsis had no influence on emotion ratings, suggesting that whilst all of the devices studied can influence how ironic a comment is perceived to be; only emoticons have an effect on the emotional impact of the message.
Eye-tracking evidence shows that people prefer to inspect the target of a recently-depicted event vs. that of a possible future event while comprehending a spoken sentence describing the recent or future event. Moreover, this preference appears to persist even when during the experiment future events are shown and referenced to more often than past events.

In two visual world studies (N=32 each), we tested the recent event preference against a situational cue, which has been found to be highly effective in guiding visual attention, i.e. gaze (Hanna & Brennan, 2007; Knoeferle & Kreysa, 2012). Participants saw the experimenter sitting at a table with two objects on either side (e.g., pancakes/strawberries). After seeing the experimenter perform an event (sugaring strawberries) participants listened to an NP1-VERB-ADV-NP2 German sentence, in the past (lit. ‘The experimenter sugared recently the strawberries’) or in the present tense with a future meaning (‘The experimenter sugars next the pancakes’). During the verb (Expt 1) or at verb onset (Expt 2) the experimenter either shifted the gaze towards the object that would be named in the sentence (NP2), or looked straight ahead. After the sentence participants saw the experimenter performed the ‘future’ action. Thus the design was 2 (object: recent vs. future) x 2 (tense: past vs. future) x 2 (gaze: gaze vs. no-gaze).

We analyzed participants’ eye fixations to the two objects in the display (pancakes/strawberries) while processing the sentence using mean log-ratio gaze probabilities \( \ln(p(\text{recent object})/p(\text{future object})) \). Fig. 1 plots these probabilities in 20ms slots from Verb onset for Expt 2 (Critical regions: VERB-ADV-NP2). Fig. 1 shows that gaze affected in particular looks to the recent object in the early VERB region. ANOVAs on mean log-ratios for each region gave similar results for both experiments, i.e. a tense effect in all regions and a gaze effect in the VERB region. However, crucially, in the VERB region there was a gaze by tense interaction: contrasts showed that gaze (vs. no-gaze) enhanced looks to the future object \((p<.001)\), but not to the recent object. Importantly, we observed an overall preference for the recent object up to the last region irrespective of gaze and tense. Therefore, as expected, gaze acted rapidly in directing attention to the future object, thus mitigating the recent event preference. However, as in previous research, there was still an overall recent object preference until sentence end. This was corroborated by results of a post-experiment gated memory test (Expt 2), which showed a trend for better memory for recent objects but no effect of gaze. We propose that what underlies the recent event preference is a deeply ingrained epistemic bias of the human mind that favors assertions about past events over future ones.

Fig. 1 Mean log gaze probability ratios (\( \ln(p(\text{recent target})/p(\text{future target})) \)) as a function of conditions from Verb onset for Experiments 1 and 2

![Fig. 1 Mean log gaze probability ratios](image-url)
KNOWLEDGE OF WHAT OTHERS KNOW CONSTRAINS DISCOURSE PROCESSING
Yuki Kamide & Anuenue Kukona (University of Dundee)
y.kamide@dundee.ac.uk

Language comprehenders’ understanding of unfolding linguistic input is rapidly constrained by information from various sources. Among these is real-world knowledge, which comprehenders use to anticipate linguistic outcomes. For example, Kamide et al. (2003) found that listeners hearing “The dog will drink…,” while viewing Fig. 1, anticipatorily fixated the water (lake) based on real-world knowledge (i.e., that dogs typically drink water, not brandy). However, an ongoing debate concerns the time course over which comprehenders’ knowledge of other people’s knowledge (i.e., perspective or Theory of Mind information) influences language comprehension. While delayed accounts assume that comprehenders process language egocentrically first, then integrate other people’s perspectives (Keysar, et al., 2000), immediate accounts assume that comprehenders integrate other people’s perspectives from the outset (Hannah et al., 2003). The aim of the current study was to distinguish between delayed and immediate accounts by examining comprehenders’ sensitivity to another character’s perspective in a short discourse, and its influence on comprehenders’ anticipatory eye movements. Critically, we investigated how comprehenders balance real-world knowledge with perspective information in a non-interactive context.

Table 1. Example discourses

<table>
<thead>
<tr>
<th>Discourse</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Josephine knows that the dog likes alcohol.</td>
</tr>
<tr>
<td>1b</td>
<td>Josephine thinks that the dog will drink the brandy.</td>
</tr>
<tr>
<td>2a</td>
<td>Josephine doesn’t know that the dog likes alcohol.</td>
</tr>
<tr>
<td>2b</td>
<td>Josephine thinks that the dog will drink the water.</td>
</tr>
</tbody>
</table>

Participants heard discourses like (1) or (2) while they viewed visual scenes like Fig 1 (preceded by another display with, e.g., a girl, and a spoken sentence, e.g., “This is Josephine”). We asked: what do participants anticipate that the dog will drink in (1b) and (2b)? In (1), both participants and Josephine have information that the dog likes alcohol. In (2), participants have this information, but Josephine does not. According to delayed accounts, participants will (initially) use their own knowledge in (2) to anticipate the brandy; according to immediate accounts, they will use Josephine’s perspective (and real-world knowledge of dogs) to anticipate the water.

During ‘drink the’, participants fixated the brandy more in (1) than (2), and they fixated the water more in (2) than (1), consistent with immediate accounts. But, while they fixated the brandy more than the water in (1), fixations to the brandy and water did not differ in (2). These findings suggest that comprehenders’ knowledge of other people’s knowledge modulates anticipatory eye movements to likely upcoming referents (e.g., see looks to the brandy in Fig. 2A vs. B). But, the considerable looking to the brandy in (2) suggests that egocentric information may still compete with this perspective information (e.g., see constraint-based approaches; Brown-Schmidt & Hanna, 2011), especially in non-interactive contexts. Currently, we are running a second experiment that examines verb selectional restrictions, which we expect to be more constraining than real-world knowledge; preliminary analyses suggest that the use of perspective may be even weaker in such contexts.

Fig. 1. Example scene

Fig. 2. Proportions of fixations to the water/brandy
SPATIAL INFLUENCES ON OBJECT ACCESSIBILITY IN DISCOURSE PROCESSING
Glenn Williams (University of Dundee*), Anuenue Kukona*, & Yuki Kamide*.
g.w.williams@dundee.ac.uk

Language comprehenders rapidly update their mental representations of events in order to accommodate changes in object location. For example, comprehenders hearing narratives such as (1), while viewing scenes such as Fig. 1, fixate the table when the book is later mentioned (Altmann & Kamide, 2009). Radvansky and Copeland (2006) argue that the structuring of event representations is dependent on the nature of the space traversed. They showed that when objects move from one room to another, participants are slower and less accurate to respond to recognition probes for the objects, compared to when objects move within one room. Radvansky and Zacks (2011) state that in the former case, objects are represented across two event models, whereas in the latter, within one model. Interference is thought to occur when we reactivate multiple event models in order to access information associated with them. Here, we explore how crossing from one room to another – or an “event boundary” – influences the updating of event representations, and the deployment of attention in the visual world paradigm. We predicted decreased accessibility for objects that cross an event boundary, and thus, fewer looks to both locations (i.e., book/table).

Participants heard (1) while viewing Fig. 1a or 1b. Scenes depicted several objects, including the book and table. Critically, the doorway separating the two rooms in the scenes was located either between (Fig. 1b) the book and table, or not (Fig. 1a). Thus, the book moved either between two rooms (between; Fig. 1b) or within one room (within; Fig 1a). During the discourse final “book,” we found that the table was fixated more in the within vs. between condition. The book was also fixated more than the table, although we found no conditional differences in fixations on the book.

Consistent with Radvansky and Copeland (2006), our results suggest that crossing from one room to another – or an “event boundary” – influences the structuring of comprehenders’ event representations. Crucially, our results are novel in that they reveal that when objects move from one room to another, participants fixate their updated locations (e.g. table) less, compared to when objects move within one room. Yet, this decrease in accessibility for the updated location of the book (on the table) when it crossed an event boundary was not associated with a change in accessibility for its initial location. Rather, the book was fixated equally across conditions. This could indicate that event boundary effects stem specifically from representations of updated locations, which is not made clear through the use of recognition probes (e.g., Radvansky & Copeland, 2006). Alternatively, it is possible that low-level processing of linguistic and visual bottom-up input (i.e., looking at the book when ‘book’ is heard) may have masked any difference in accessibility for the initial locations.

We are currently collecting data for a second experiment that addresses this issue. In order to ensure that participants’ fixations are driven primarily by their mental representations of the events (e.g., vs. low-level bottom-up processing), we have replaced the book with a bag in our scenes, and we have modified the first sentence in our narratives (e.g., “The woman will move the book from the bag to the table”). If crossing from one room to another decreases accessibility overall (e.g., Radvansky & Copeland, 2006), we expect fewer looks to both the updated (table) and initial (bag) locations. Nevertheless, our current results indicate that comprehenders’ event representations, and their attention, are modulated by the nature of the space that is traversed in an event. We discuss these results in relation to dissimilarity-based interference in event comprehension (Hindy et al., 2013).

(1) “The woman will take the book to the table. Then she will study the painting and pick up the book.”

Fig. 1. Example visual scenes for the (a) within and (b) between room conditions.
DOES PRONOUN PROCESSING VARY ACROSS LANGUAGES?
Sol Lago (U of Maryland), Shayne Sloggett (UMass), Zoe Schlueter (U of Maryland), Wing yee Chow (BCBL), & Colin Phillips (U of Maryland)
sollago@umd.edu

Pronouns carry little meaning of their own and require information about their antecedents in order to be interpreted. But what is the nature of the antecedent representations retrieved during comprehension? Previous cross-modal and production studies have shown that pronoun resolution is sensitive to different types of information [1,2,3], but findings from different tasks and languages suggest that antecedent representations might differ across languages: specifically, both discourse and lexical-level antecedent representations are indispensable in languages with grammatical gender such as German, Spanish and Italian, which store noun gender information in the lexicon [2,3]. In contrast, only discourse-level information may be needed in languages that lack grammatical gender such as English [4,5]. To test this generalization, in two eye-tracking studies we examined whether semantic and/or phonological antecedent information was reaccessed during pronoun processing in German and English. We found that German, but not English speakers, show semantic priming effects during coreference. We propose that speakers of languages with and without syntactic gender access different types of antecedent representations during comprehension.

In the present experiments (n = 60 each) we varied a pronoun’s antecedent such that it was either related (singer) or unrelated (deputy) to the word immediately following the pronoun. This relationship was either semantic (piano) or phonological/orthographic (sink). We predicted that upon encountering the possessive pronoun, comprehenders should retrieve some representation of its antecedent, and that the antecedent properties retrieved should impact reading times to the word after the pronoun. To isolate relatedness effects specifically due to referential processing, we manipulated whether a pronoun or a determiner preceded the target word, resulting in a 2 (related/unrelated) × 2 (semantic/phonological) × 2 (pronoun/determiner) design. We found that German comprehenders showed immediate facilitation effects on the word after the pronoun (significantly shorter reading times in single fixation, first fixation and first pass reading times) when the word and the pronoun’s antecedent shared a semantic relationship. These effects were specific to the pronoun conditions, as shown by a relatedness × determiner type interaction. In contrast, whereas English comprehenders showed facilitation in the semantic conditions, this effect was shared across pronouns and determiners, consistent with the claim that it was likely due to priming from the entire sentence context, and not to antecedent reactivation due to coreference. The absence of semantic priming effects in English suggests that coreference in this language relies only on discourse-level information, and that semantic priming effects might be tied to retrieval of syntactic gender information.

**Semantic conditions**
Related/Unrelated. The maintenance men told the singer/deputy about a problem. They had broken the/his piano and would have to repair that first.

**Phonological conditions**
Related/Unrelated. The maintenance men told the singer/deputy there would be a delay. They said that the/his sink wouldn't be installed until next month.

Sentence comprehension may depend substantially not only on the form of utterances, but also on comprehenders’ CAUSAL MODELS for why speakers choose particular forms. We show that an appeal to such models adjudicates between two contrasting views of the representation of syntactic alternations: (i) that distinct forms have distinct meanings, and (ii) that distinct forms are synonymous, and chosen for grammatical or processing reasons. In particular, comprehenders infer an interaction between both factors—intended meaning and grammatical probability—in their causal models of speaker production.

The dative alternation comprises two syntactic structures with (potentially) equivalent meanings, the prepositional object (PO) structure (1a) and the double object (DO) structure (1b). On a construction grammar view, speakers choose POs to convey change of location, and DOs to convey change of possession (Goldberg, 2006; Gries et al, 2005; Kaschak & Glenberg, 2000). However, no study directly investigates comprehender inferences of these contrasting meanings. In contrast, a gradient grammar view holds that the structures have overlapping meanings, and speakers choose the structure that minimizes processing cost by, for example, placing short constituents before long (Bresnan et al, 2007; Ferreira, 1996).

We investigate whether comprehenders infer that both intended meaning and processing considerations interact in speaker choice using an alien language game paradigm. To probe comprehender inferences of constructional meaning, we constructed sentences with nonsense words and asked participants to choose between possessive and locative inferences (below). We crossed SYNTACTIC STRUCTURE with GRAMMATICAL PROBABILITY by jointly manipulating length and definiteness (Bresnan et al, 2007), so that grammatically likely sentences had short, definite constituents before long, indefinite ones (opposite for grammatically unlikely items). Each of 60 Mechanical Turk participants saw 5 of 100 critical items (e.g. [1a] below), each with a different inference probe, plus 15 fillers.

Construction grammar accounts predict a main effect of syntactic structure: comprehenders should make more locative inferences for PO structures. Further, if comprehenders infer that both intended meaning and processing considerations interact in speaker choice, we predict that locative inferences will be especially frequent for grammatically unlikely POs. In other words, comprehenders should infer that speakers choose a structure in spite of low grammatical probability only when they especially strongly intend the meaning associated with that structure. A maximal mixed-effects regression revealed both the predicted main effect ($p < 0.0001$) and interaction ($p < 0.01$). Subsequent followup experiments separately manipulating constituent length and definiteness suggest that of these grammatical factors, length is the primary driver of the interaction.

We thus demonstrate for the first time that comprehenders infer interaction between two determinants of their interlocutors' syntactic choice during production: meaning intent as predicted by construction grammar accounts, and grammatical probability as predicted by gradient grammar accounts. This result suggests that comprehenders build fine-grained causal models of their interlocutors' productions, which motivates future research on the extent to which syntactic representations in online comprehension comprise a joint probability distribution over inferred production considerations such as intended meaning and grammatical processing factors. Future research should also explore the extent to which these inferences are informed by comprehenders' own production preferences.

(1) a. The zarg prolted the cherid to a really gromious flig.  
   b. The zarg prolted the flig a really gromious cherid.

(2) a. The zarg prolted a really gromious cherid to the flig.  
   b. The zarg prolted a really gromious flig the cherid.

Which is more likely?
— The cherid has a new owner.  
— The cherid was moved to a new place.
THE PROCESSING OF MORPHOLOGICALLY COMPLEX WORDS IN TURKISH HERITAGE SPEAKERS

Gunnar Jacob (Potsdam Research Institute for Multilingualism, University of Potsdam) & Bilal Kırkıcı (Middle East Technical University)

gujacob@uni-potsdam.de

Previous experimental studies on the processing of morphologically complex words have mainly focused on either native speakers or late L2 learners. In the present study, we investigate morphological processing in so-called heritage speakers who have acquired a minority language as their first language (L1). Although heritage speakers are native speakers in the sense that they acquired their L1 from birth, some previous studies have reported properties characteristic of L2 learners, particularly with regard to morpho-syntactic phenomena (e.g. Montrul, 2002; Rothman, 2007).

The current study is (to our knowledge) the first to investigate morphological processing in heritage speakers. In two masked-priming experiments, we compared the processing of morphologically complex Turkish words in heritage speakers of Turkish raised and living in Germany to two groups of non-heritage speakers from Turkey, a group of L1 native speakers of Turkish (T-L1) and a group of L2 learners of Turkish (T-L2). The materials and the data from the T-L1/T-L2 groups were taken from Kırkıcı and Clahsen (2013). The design was also held parallel, to allow for direct comparisons between the three groups. Experiment 1 investigated priming effects for morphologically related prime-target pairs. Heritage speakers showed the same priming effects as the T-L1 group, with significant priming effects for prime-target pairs with inflected primes (e.g. ‘sorar-sor’ asks-ask) as well as for prime-target pairs with derived primes (e.g. ‘sağlık-sağ’ health-healthy). In Experiment 2, we measured priming effects for prime-target pairs which were semantically and morphologically unrelated, but only related with regard to formal overlap (e.g. ‘devre-dev’ period-giant). Interestingly, unlike both the T-L1 and the T-L2 groups, heritage speakers also showed significant priming effects in this condition. Importantly, the priming effects in Experiment 2 were not weaker than in Experiment 1. As formal overlap alone was enough to cause priming effects, the effects heritage speakers showed in Experiment 1 are probably not morphological in nature, but instead due to the formal (orthographic) overlap between prime and target.

Our results suggest that heritage speakers differ from both native speakers and L2 learners in that they rely more on surface form properties of the stimulus (viz. orthography) during early stages of word recognition.

References


This study investigates effects of word predictability in reading. The question is whether older adults benefit from local predictability to the same extent as do younger adults and whether predictability differentially affects silent reading vs. oral reading, the latter involving an additional speech planning and articulation stage.

One way to operationalize predictability is Transitional Probability (TP) which indicates how likely a word is conditioned by its left or right neighboring word. TPs are computed via frequency counts from large language corpora and reflect probabilistic knowledge on how likely a particular word is to come next. Several reading studies (e.g., McDonald & Shillcock, 2003, *PsychScience*) have shown that frequent word combinations (such as ‘to give advice’) can be processed faster than less frequent combinations (such as ‘to give ideas’). Furthermore, frequent and predictable words are articulated with less effort in speech production (e.g., Bell et al., 2009, *JML*). Given that frequency and predictability effects are based on language experience, they are prime candidates for age-related changes. However, evidence is mixed on whether frequency and predictability effects generally decrease or increase with age in adulthood. In a corpus study on read aloud speech we found that older adults show slightly smaller-sized (TP) predictability benefits in their spoken word durations than young adolescents. Predictability in reading aloud may exert its effect either on the comprehension level, or the speech planning and pronunciation level, or both. The aim of the present study was to separate these various levels by studying whether TP affects eye fixation behavior differently in silent and oral reading, and to compare TP effects on eye fixations to TP effects on spoken word duration data.

In the present study 30 younger and 30 older participants read full sentences both silently and aloud. These 240 sentences included critical verb-noun combinations in which transitional probabilities and word frequencies varied continuously from high to low values. The analysis of the fixation durations shows clear predictability (TP) effects, that were of equal size in oral vs. silent reading. Second, the fixation data do not show a predictability by age group interaction, indicating that older and younger adults predict upcoming words to the same extent in both reading modalities. We further analyze skipping rates and refixation rates for the target verbs. Older adults have been shown to employ a different reading strategy than younger adults with higher skipping rates, but also higher refixation rates (Rayner et. al., 2006, *Psych&Aging*). Our data confirm these effects, with older adults skipping target verbs more frequently while at the same time regressing to the verb area more often than younger adults. High-predictable verbs are skipped more often than low-predictable words, especially in the younger adults’ sample. However, skipping occurs rarely (less than 6 % of the total sample) and the age by predictability interaction in our sample is thus a marginal effect. Analysis of the spoken word durations from the oral reading condition is in progress and will be discussed.

We conclude that local predictability influences reading behavior and speech planning. Older adults use probabilistic knowledge as much as younger adults in order to process immediately upcoming words. Our data thus differ from findings of age-related differences in contextual prediction reported in the aging literature (e.g., Federmeier et al., 2010, *Brain&Lang*). This difference in results may relate to whether predictability effects are local (as in TPs), or whether the cues that predictions are based on need to be integrated over a sentence (as for instance in cloze tasks).
THE EFFECT OF EXPOSURE IN CROSS-DIALECT PERCEPTION: HEARING AMBIGUOUS /r/ VARIANTS IN GLASWEGIAN

Robert Lennon (University of Glasgow).
rlennon.1@research.gla.ac.uk

In Glasgow, speakers are stereotypically rhotic. However, recent sociophonetic research indicates a trend towards the loss of postvocalic /r/ in working class (WC) Glaswegian speech, leading to ‘derhoticisation’ (Stuart-Smith 2007), while a simultaneous increase in rhoticity in middle class (MC) Glaswegian speech is underway (Lawson et al. 2011). This means there is potential for misperception when listeners hear minimal pairs such as e.g. hut/hurt, when spoken by WC derhoticising speakers (who realise postvocalic /r/ as a pharyngealised variant), but less so when the same words are spoken by hyper-rhotic MC speakers (who realise /r/ as a bunched-tongue ‘schwar’ – a strongly rhotic vowel with a lowered third formant). It is hypothesised that the pharyngealized nature of derhoticised /r/ makes it perceptually very similar to the preceding open back vowel in /CʌC/, leading to difficulty when listeners try to distinguish between the two segments. This difficulty is hypothesised to decrease as experience of the Glaswegian linguistic environment increases (Adank et al. 2009).

While we know something about the acoustic, auditory and articulatory features of derhoticised /r/ (e.g. Lawson et al. 2011), this paper presents the first research to investigate how the variant is perceived by different listeners, and how the amount of exposure to the Glaswegian accent can affect spoken word recognition. A perceptual experiment investigated whether a listener’s amount of exposure to the accent affects their ability to identify the word the speaker intended to produce, given a choice from words in minimal pairs with and without postvocalic /r/.

62 participants were from three accent groups, with varying levels of experience of Glaswegian:

1. Raised in Glasgow, living in Glasgow;
2. Raised in England, living in Glasgow, attending Glasgow University (mean = 3.6 years);

In a two-alternative-forced-choice (2AFC) task, all participants were asked to choose which word they thought they heard, out of minimal pairs such as hut/hurt, spoken by both derhoticising and hyper-rhotic speakers. In a separate task, they rated the ‘strength’ of the /r/ sound in each stimulus. Stimuli were 192 words segmented from continuous speech, elicited from both WC and MC native Glaswegian speakers carrying out a structured collaborative task. The data were analysed using linear mixed effects modelling in R.

Results from the 2AFC task showed that the presence of WC derhoticised /r/ caused perceptual ambiguity for all listener groups, with the MC schwär variant eliciting much more accurate responses. Nonetheless, Glaswegian listeners were the most accurate and the fastest, while English listeners in Cambridge were by far the least accurate and the slowest. English listeners in Glasgow displayed an intermediate, complex pattern, which shows an intriguing effect of longer-term learning of derhoticised /r/). In the rating task, MC schwär elicited a strong /r/ rating from all listener groups, following predictions. However, there was more variation in responses to the derhoticised /r/ stimuli, with listeners living in Cambridge giving much weaker ratings for derhoticised /r/ tokens than listeners in both Glasgow groups. These results provide support for the hypothesis that familiarity with an accent’s fine-grained phonetic detail aids word recognition. Specifically, the Scottish listeners are the most efficient at recognising derhoticised variants as /r/, English listeners living in Glasgow show a pattern which is almost as efficient, and listeners living in Cambridge find it the hardest to identify derhoticised tokens as /r/-ful.
DIRECTIONALITY AND UNDERSPECIFICATION: THE PROCESSING OF ZERO-DERIVATION

Jeannique Darby (University of Oxford) jeannique.darby@ling-phl.ox.ac.uk

It has been argued that forms in English which are used as both nouns and verbs (e.g., CLOAK, BITE) may be classified according to which form is derived from the other, in a process frequently called 'zero-derivation' or 'conversion' (cf. Marchand, 1963; Kiparsky, 1982; Beard, 1995). A further claim is that, as in derivation via affixation, an item derived through conversion is more 'morphologically complex' than its base. For example,

a) Verb Base (V_B): (to) [BITE_v] less complex than (a) [[BITE_v]N]

b) Noun Base (N_B): (a) [CLOAK_N] less complex than (to) [[CLOAK_N]V]

An alternative view argues that there is no difference in complexity, and that directionality is not synchronically represented in the morphology (e.g., Lieber, 1981; Farrell, 2001). Thus, in some cases, directionality is not always clear: e.g., is the verb GUARD derived from the noun ([[GUARD_N]V]: ‘to act as a guard’), or is the noun derived from the verb ([[GUARD_V]N]: ‘one who guards’)? Farrell (2001) suggests that neither form is derived, and that such ‘pairs’ may in fact be a single entry in the lexicon, underspecified for word class.

To experimentally examine these opposing views, two lexical decision tasks were conducted using the delayed priming paradigm, which previous research (Henderson et al., 1984; Drews and Zwitserlood, 1995) has suggested may reveal effects of morphological processing that are independent of the effects of purely semantic or phonological relatedness. Experiment 1 examined the relative effect of prior presentation of a related -ing form on four different types of monosyllabic target words: basic nouns (N_B Nouns); verbs derived from these nouns (N_B Verbs); basic verbs (V_B Verbs); and their converted nominal counterparts (V_B Nouns). Experiment 2 investigated the priming of nouns that were members of directional pairs (N_B Nouns and V_B Nouns) as compared to frequency-matched non-directional items (uN_B Nouns and uV_B Nouns). Directional items were grouped based on surveys of speaker intuition; the underspecified items (judged by speakers as ‘non-directional’) were further categorised based on CELEX tags (Baayen, et al., 1993), which were also matched to speaker ratings for directional items. Each of the target groups in the two experiments appeared 6-8 items after either the -ing form of the verb (e.g., cloaking), or an unrelated -ing form. The results are illustrated below (Mean Unrelated response time minus Mean Related response time in ms; * = significant at p<.05; ** = significant at p<.01).

Experiment 1 showed an asymmetry in the priming of V_B items and N_B items. Since the primes are verbal forms, both types of related -ing activated their respective verbs. However, the additional effort required to further decompose an N_B prime like [[[cloak_N]V]ing] into its base noun mitigated the facilitation of the N_B Noun, leading to a small effect which was not statistically significant. Experiment 2 then revealed a greater priming effect for non-directional nouns than for directional ones. This suggests that, despite the difference in the underlying morphological structure of N_B and V_B -ing primes, they are both less closely related to their respective nouns than the non-directional primes. Together, the results support the view that directionality in some conversion pairs is represented on a morphological level in the grammar, but that some pairs may not only be non-directional, but may even be a single lexical entry without a fixed specification of word class.
The memory retrieval mechanism employed by the human sentence processor in dependency resolution is, at least in some cases, susceptible to interference from items that are in a structurally inaccessible position but partially match the retrieval cues. In English, interference effects are seen in subject-verb agreement and occasionally in reflexive anaphors [1,2,3,4]. The cue-based retrieval mechanism commonly used to explain interference effects [5] predicts (A) INHIBITION for ACCESSIBLE-MATCH INTERFERENCE: A slow-down in retrieval latency when both the structurally accessible and the inaccessible antecedent match, e.g., the gender cue, and therefore compete for a limited amount of spreading activation from that cue; and it predicts (B) FACILITATION for ACCESSIBLE-MISMATCH INTERFERENCE: When the accessible antecedent is not a full match there is a speed-up due to occasional, unnoticed misretrieval of a partially matching interfferer [cf. 6].

In subject-verb agreement, the observed interference effects match the predictions [3,4]. However, for argument reflexives, prediction (B) is controversial: Among the studies that found an effect, [7] confirms a facilitation for ACC-MISMATCH INT. and others support inhibition [8] or show signs of both in different measures or individuals [2,6,9]. A clear support for INHIBITORY interference was found in a study by [10] on Mandarin reflexives, where the reflexive ZIJI requires a c-commanding, animate antecedent: A slow-down was found in gaze durations and right-bounded reading time in the ACC-MISMATCH condition when the inaccessible antecedent matched the animacy cue vs. when it did not. Inhibition in ACC-MISMATCH INT. cannot be explained by the current assumptions of the cue-based retrieval account, as the two antecedents do not overlap in their match to the retrieval cues. We show that if cues are confusable, the cue-based model predicts that inhibitory interference occurs. In Mandarin, if readers treat the +C-COMMAND and +ANIMATE cues as similar, this makes the two potential antecedents compete for activation, increasing mean retrieval latency.

We present results from computational modeling in ACT-R that show: (1) with perfectly distinguishable cues, ACC-MISMATCH INT. cannot be predicted as inhibitory; (2) with increased cue similarity, the model predicts a slow-down as observed in Mandarin reflexives; (3) variable amount of cue similarity leads to inhibition, facilitation, or absence of an effect. With a similarity range of -1 (dissimilar) to 0 (identical), the Mandarin gaze duration data is best fit with a value of -0.6. Values around 0.8 predict no effect; lower and higher values predict increasing facilitation or inhibition, respectively. Consistent with the evidence, however, the quality of the inhibition effect in the ACC-MATCH condition is not affected by cue similarity.

Our modeling results suggest that in the case of reflexives, Chinese readers do not perfectly distinguish structural from non-structural cues. Furthermore, we show that qualitatively different experimental results could be explained as emerging from variable cue similarity. The proposal therefore implies that, just like items in memory, even their features can be confusable and this might be modulated not only by language but also by task and individual differences. If keeping cues separated requires cognitive effort, it would mean that readers dynamically adapt cue similarity to task-requirements. The unequivocal evidence for facilitatory interference in subject-verb agreement compared to reflexives could thus be the result of a task-specific setting of similarities. Differential interference patterns for high-span and low-span readers in [9] speak in favor of individual differences in cue separation. In sum, our proposal offers a principled explanation for the equivocal findings concerning interference in reflexive anaphors by extending the cue-based retrieval account with cue confusability.

In John said Bill disliked himself/him, the Binding Theory (BT) (Chomsky, 1981) regulates that the reflexive ('himself') must be bound by the local subject ('Bill'), but not by the matrix subject ('John'), whereas the pronoun ('him') can be co-indexed with the matrix subject but not with the local subject. Existing work on processing of English reflexives and pronouns has yielded mixed results regarding 1) whether the early stage of processing is guided by structural constraints alone (Sturt, 2003; Xiang et al., 2009) or also affected by non-syntactic factors (e.g. gender) (Badecker & Straub, 2002) and 2) whether different referential types exhibit similar (Nicol & Swinney, 1989) or asymmetrical sensitivities (Kaiser et al., 2009) to multiple constraints. To further explore these issues, we examined, using self-paced reading, Chinese complex reflexives ta-ziji (s/he-self) – a real counterpart to English reflexives but has largely been neglected (cf. nearly exclusive literature on the bare reflexive or logophor ziji), and contrasted its processing with that of Chinese pronouns ta (s/he).

Exp. 1 (N=54) focused on the reflexive. Depending on the gender of ta-ziji (s/he-self), we manipulated the gender of i) the (BT-)accessible NP 2 (match/mismatch) and ii) the (BT-)inaccessible NP 1 (match/mismatch), yielding four conditions (1). At ta-ziji, we found a marginal main effect of the gender of Inaccessible NP 1 (β=-0.07, SE=0.04, t=-1.82): Gender-matching conditions were read slower than mismatching ones, indicating immediate interference from gender-matching but structurally inaccessible NPs. At the 1st spillover region ('and'), we found a main effect of the gender of Accessible NP 2 (β=0.11, SE=0.03, t=3.24), but no other effects: Gender-matching conditions were read faster than mismatching ones, as predicted by the BT principle A (BT-A).

With the same design, Exp. 2 (N=54) focused on the pronoun ta (s/he), but now NP1 was accessible, and NP2 inaccessible. No effects were found at ta. In the 1st ('and') and 3rd ('DE') spillover regions, we found a main effect of the gender of Accessible NP 1 (β=0.10, SE=0.03, t=3.26; β=0.06, SE=0.02, t=2.72): Gender-matching conditions were read faster than gender-mismatching conditions. In the next region ('other'), we found a main effect of the gender of Accessible NP 1 (β=0.06, SE=0.02, t=2.62) and a marginal interaction between the gender of NP 1 and NP 2 (β=-0.06, SE=0.03, t=-1.85): the condition with accessible mismatching NP 1 and inaccessible matching NP 2 was read slowest, suggesting that only when accessible NPs mismatch the pronoun in gender could the gender of inaccessible NPs exert interference effects.

Our findings demonstrate that non-structural gender could exert interference effects in the processing of Chinese pronouns and complex reflexives. However, the interference effects differ in accordance with different referential forms. While ta-ziji demonstrates early sensitivity to non-structural cues, thereby resembling ziji (Chen et al., 2012), the processing of ta is affected at a later stage by non-structural gender than that of complex reflexives.

1) a/b: NP1-match/mismatch, NP2-match ; c/d: NP1-match/mismatch, NP2-mismatch
NP1[say][NP2][repeatedly][remind][s/he-(self)][and][class][DE][other][examinee][bring][all][DE][certificate].

‘NP1 says that NP2 repeatedly reminds him/her-self and other examinees in the class to bring all the certificates’.
The extent to which filler-gap dependency formation in non-native (L2) compared to native (L1) language processing may differ has received considerable attention in the L2 processing literature [1]. Some studies reported that non-native speakers show immediate sensitivity to filled-gap effects but delayed evaluation of a filler’s thematic fit, particularly in tasks which do not require overt evaluation of plausibility [4,5]. Others have reported immediate sensitivity to plausibility manipulations but delayed filled-gap effects [2], which may suggest gap-filling in a non-native language is initially guided by semantic goodness-of-fit evaluation rather than constituent structure information. However, in [4,5] filled-gap and plausibility effects were manipulated within the same sentence, which could potentially mask effects, whereas [2] based their conclusions on a between experiment comparison.

We conducted an eye-movement experiment which manipulated both plausibility (1a,b) and filled gaps (2a,b) in the same experiment. 32 mature native English speakers and 32 native German speakers of L2 English read 24 sentences as in (1,2), plus 60 fillers, while their eye movements were monitored. For (1), reading times at had written should be longer for implausible condition (1b) than plausible (1a). Reanalysis in (1) at the ultimate gap site ‘about’ may also give rise to a reverse plausibility effect [3], indicative of the relative ease of revising an initially implausible dependency. In (2), longer reading times at ‘documents’ in (2b) than (2a) would be diagnostic of a filled-gap effect.

(1a) Fred liked the book that the journalist had written quickly and very passionately about at some point earlier in the afternoon.
(1b) Fred liked the city that the journalist had written quickly and very passionately about at some point earlier in the afternoon.
(2a) Fred liked the book about which the journalist had written documents very passionately at some point earlier in the afternoon.
(2b) Fred liked the book that the journalist had written documents very passionately about at some point earlier in the afternoon.

For the plausibility manipulation, at the critical and spillover regions (‘had written’ and ‘quickly and’) we observed significant main effects of plausibility that were not modulated by language background. In first- and second-pass measures, L1 and L2 speakers had longer reading times in implausible condition (1b) than plausible condition (1a). At the ultimate gap site (‘about’), where the initial dependency needs to be revised, reading times tended to be shorter for implausible condition (1b) than (1a) for both groups, indicating that an initially implausible dependency was easier to reanalyse. For filled-gaps, at the critical and spillover regions (‘documents’ and ‘very passionately’ respectively) we observed significantly longer reading times in filled-gap condition (2b) than (2a) for both L1 and L2 speakers in both first- and second-pass measures. In second-pass measures, this was modulated by a significant interaction with language background. Although both groups showed reliable filled-gap effects, the filled-gap effect was larger for the L2 compared to the L1 group.

In summary, we observed no evidence of a temporal delay in sensitivity to either goodness-of-fit or filled gaps for non-native compared to native English speakers. Instead, both groups showed a similar time-course for both types of diagnostic for dependency formation, suggesting that the processing of direct object gaps is qualitatively similar in L1 and L2 sentence processing.

References
WH-DEPENDENCY RESOLUTION INTERRUPTS REFLEXIVE ANTECEDENT SEARCH ON-LINE

Michael Frazier, Peter Baumann, Lauren Ackerman, David Potter & Masaya Yoshida
(Northwestern U)
fraze@u.northwestern.edu

In real-time sentence processing, the human parser may encounter conflicts between the parse that is most consistent with the immediately local input and the parse that is most consistent with the entire input string. "Good enough" theories of sentence processing [Ferreira et al. 2002] posit that the parser may fail to arrive at a definite interpretation of an input string and retain interpretations that, while globally impossible, fit a substring of the input, producing parses of "locally coherent" substrings [Tabor et al. 07; Konieczny et al. 09]. This study aims to determine whether the parser experiences such "good enough"/local-coherence effects when establishing reflexive pronoun dependencies that are interrupted by the tail of a long-distance wh-dependency. Interrogative embedded infinitive clauses (1a) are an ideal testing ground for crossing wh- and reflexive dependencies because, for some infinitive-complement taking verbs, omitting a sentence-initial wh-word yields a grammatical sentence (1b). Thus examples like (1a) contain a local coherence from *did* to sentence end.

1a. Which man did John expect to have injured himself?
1b. Did John expect to have injured himself?

A reflexive normally co-refers with its closest potential antecedent: *himself* = *the man* in (2a). In (2b), however, where the wh-phrase *which man* has moved from after *expect*, the wh-phrase must be the antecedent of *himself* (*himself* = *which man*), instead of the linearly closer *Jane*.

2a. Jane expected the man to have injured himself.
2b. Which man did Jane expect to have injured himself?

If the wh-dependency is resolved inside the locally coherent string, the dependency tail may interrupt the locally coherent parse, as would a normal NP, so *himself* may take the wh-word as its antecedent online. If "good enough" parsing is not sensitive to the tail of the wh-dependency, however, interference from the gender of the grammatically inaccessible but "locally coherent" antecedent should be seen.

Experiment 1 (n=24): An eye-tracking text reading experiment with a 2x2 design manipulated the gender match between the reflexive and the grammatically acceptable wh-phrase antecedent and the intervening, locally coherent but globally unacceptable subject antecedent. In the paradigm (3), we investigated whether wh-dependency resolution interrupts the potential local coherence. If so, the wh-phrase should be chosen on-line as the antecedent of the reflexive in (3). If the locally coherent parse is chosen, the closest overt NP (Anna/Steven) should be initially interpreted as the antecedent. Which antecedent the parser chooses can be measured via gender-mismatch effects [e.g. Sturt 2003].

(3) Which cowgirl did Anna/Steven expect to have injured himself/herself … ?

Analysis of regression path time reveals significant results. There was a main effect of the accessible (wh-phrase) antecedent's gender congruency with the reflexive (ps<.05). That is, the reflexive region was read significantly slower in Wh-phrase Mismatch condition than in the Wh-phrase Match condition. On the other hand the grammatically inaccessible subject NP (Anna/Steven in 2) did not affect the reading of the reflexive (regression path; ps<1).

Experiment 2 (n=20): Experiment 1 has an alternate explanation whereby subjects adopt an object-extraction reading (e.g. Which man did Jane expect PRO to have injured t) such that the slowdown is due to PRO/trace competition (Frazier et al. 1983). A follow-up experiment used finite complement clauses which do not admit of this reading and contain no potential "good enough" string. Preliminary results indicate that the reflexive region was read significantly slower in Wh-Mismatch condition than in Wh-Match condition, as in Exp. 1.

(4) Which cowgirl did Anna/Steven expect had injured himself/herself … ?

These results indicate that the parser selects the linearly further but globally grammatical wh-phrase as the antecedent of the reflexive in online reading, rather than the intervening subject NP which is a potential locally coherent antecedent for reflexive binding. This result strongly suggests that the resolution of a wh-dependency suffices to interrupt a potential local coherence and permit the parser to build globally grammatical structure in these contexts immediately, indicating that parsing of these kinds of crossing dependences is highly grammatically constrained, and in this sense more than "good enough".
AGREEMENT INHIBITS DETECTION OF SPRAY/LOAD MISMATCHES IN GAPPING
Michael Frazier & Masaya Yoshida (Northwestern U)
fraze@u.northwestern.edu

Prior research [Frazier & Yoshida 2013] suggests that mismatch between the agreement form a Gapped verb would possess and the overt agreement of its antecedent induces difficulty effects after the disambiguation region, indicating that the Gapped clause’s antecedent may be retrieved by feature-matching of a morphosyntactic representation from a content-addressable memory store [Lewis & Vasishth, 2005; Lewis et al., 2006], as Wagers et al. (2009) propose for similar interference effects in S+V agreement.

In contrast to agreement mismatch, argument-structure mismatch, such as that between the two verb phrase configurations of a spray/load alternative verb load the packages in the van and load the van with packages, between a Gapped verb and its antecedent yields unacceptability. This experiment seeks to investigate whether agreement-features and argument-roles are used together to retrieve the antecedent of the missing material in the Gapped clause, or whether one pre-empts the other. On the one hand, we might expect argument-roles to pre-empt agreement-features in retrieval, since only the former are grammatically relevant in this context. On the other hand, argument-roles are not morphologically marked on the verb in English, and may for this reason be a less accessible cue for retrieval if more complex processes (e.g., plausibly, structure building of the verb phrase domain) must occur before they can be accessed.

Yoshida et al. (2013) showed that Gapping parses are incrementally preferred under parallelism with a preceding context where either a Gapping parse or a parse with a PP modifying an NP is available, as in (1a), but that a Gapping parse was not adopted when the argument structure of the potential antecedent was inappropriate for the Gapping remnants. Incorporating an agreement match/mismatch manipulation enables investigation of whether mismatched number of the subject can influence spray/load-mismatch detection.

Experiment (n=20): An eye-tracking text reading experiment used the paradigm (1), with both the agreement and the spray/load alternation cross-varied for a two-by-two design.

1a. The boys load the van with packages and the men with boxes happily wave at them.
1b. The boy loads the packages in the van and the men with boxes happily wave at him.
1c. The boys load the packages in the van and the men with boxes happily wave at them.
1d. The boy loads the van with packages and the men with boxes happily wave at him.

If agreement-features pre-empt argument-roles in antecedent-retrieval, an agreement mismatch effect should be observed at the disambiguation region (the PP argument that determines whether a Gapping parse is possible or not; in 1a, *with boxes*) in agreement-mismatched conditions as the parser predicts a Gapping construction regardless of the spray/load alignment. Conversely, if argument-roles pre-empt agreement-features in retrieval, the agreement mismatch effect should be confined to (1d), since the spray/load mismatch will have already ruled out the Gapping parse in (1b).

The spillover region (a manner adverb; in 1 *happily*) shows a significant effect of agreement match/mismatch in first-pass times (as well as a marginal effect in regression-path times) and a significant interaction between agreement match/mismatch and spray/load match/mismatch in regression-path times, such that in first-pass times, agreement-matched conditions (a,c) were read significantly slower, and in regression-path times, examples with matched agreement plus matched spray/load alternation (a) were read significantly slower (and agreement-matching conditions (a,c) marginally slower overall).

This result suggests that agreement-features are retrieved early and/or automatically in antecedent retrieval in Gapping constructions, and can inhibit the detection of grammatically-relevant argument-structure features. The interaction between agreement match and argument-structure match in regression-path times indicates that the representation retrieved is morphosyntactically rich, but the confinement of this effect to a late measure suggests that argument-structure features are pre-empted online by agreement-features. Overall, this result supports an automatic feature-matching mechanism for antecedent retrieval of unpronounced Gapped material which operates over morphosyntactically rich representations, accessing more visible features first. The direction of the effect (slower in matched conditions) is attributable to the parser’s discarding of a potential Gapping parse cued by agreement matching rather than, as in [Frazier & Yoshida 2013], having it confirmed.
SATURATING THE ANTICIPATORY MECHANISMS OF SYNTACTIC PRIMING DURING SITUATED LANGUAGE UNDERSTANDING

Eunice Fernandes (Universidade de Lisboa), Moreno I. Coco (Universidade de Lisboa) & Holly P. Branigan (University of Edinburgh)

fernandes@campus.ul.pt

Syntactic priming is the tendency to repeatedly adopt a recently processed syntactic structure across otherwise unrelated sentences (e.g., Arai et al., 2007; Bock, 1986). For example, participants in visual world paradigms show patterns of anticipatory looks that are consistent with parsing the sentence in the same way as a previous sentence. Recent research has suggested that such effects may be cumulative over the course of an experiment, based on overall relative frequency of alternative structures within the experiment (e.g., Fine et al., 2013; Kaschak, 2007). But priming effects also appear to be linked to expectations about the occurrence of a particular structure, with more surprising structures priming more strongly than less surprising ones (e.g., Jaeger and Snider, 2013). Depending on how surprisal is calculated (e.g., how contextually fine-grained it is), we might expect that the more frequently a structure has been recently experienced, the less surprising it will be. This should result in reduced anticipatory effects, and perhaps induce anti-persistence effects, whereby the alternative non-primed structure is considered. We investigated this hypothesis in a visual world eye-tracking paradigm by testing whether the consecutive repetition (i.e., the amount of short-term activation) of a particular structure would invert the anticipatory priming effect that has been observed following a single prime.

Portuguese participants (n=24) read aloud two sentences: either 2 relative clause (RC) prime sentences (Activation-2), both syntactically disambiguated for high- or low-attachment (HA vs. LA) (e.g., The relative [sg] of the boys [pl] relatives [pl] of the boy [sg] who [sg] will [sg] ...); or a filler sentence followed by a single RC prime sentence (either HA or LA) (Activation-1). They then heard a target sentence contextually disambiguated to an HA or LA interpretation (e.g., The husband of the woman who will wear the tie/ the necklace is tall while viewing a scene depicting the mentioned entities (and 2 distractor entities). This resulted into a 2 (Prime: HA/LA) x 2 (Activation: 1 vs. 2) design. We analysed proportions of fixations to the possible antecedents of who (‘Subj1’ and ‘Subj2’: the husband, the woman and the objects acted upon by those subjects (‘Obj1’ and ‘Obj2’: the tie, the necklace) from 200ms after pronoun/will onset until determiner onset of the NP following the main verb (1000 ms) aggregated over temporal bins of 50ms each. We modelled proportion of fixations as a function of prime, activation and time (orthogonal polynomial of order 2) using LMEMs.

For Activation-1 trials, we replicated previous findings: entities associated with the primed RC structure received more anticipatory looks, e.g., after Activation-1 HA primes we observed anticipatory looks to Subj1 after the onset of the pronoun who. However, for Activation-2 trials, we found reversed priming effects, i.e., anticipatory looks to entities associated with the unprimed RC structure, e.g., after Activation-2 HA primes we observed anticipatory looks to Subj2. Our data (corroborated by analyses on Obj1 and Obj2) suggests that syntactic expectations change rapidly: as soon as people are exposed to the same structure twice, they start evaluating its non-primed alternative. Accounts based only on recent or previous experience of structures do not explain the complete pattern of results.

When interpreting the meaning of a sentence, a comprehender needs to answer the basic question of who did what to whom. For some types of constructions, grammatical role assignment occurs quickly. For example, in garden-path sentences like (1) there is a period of ambiguity where it is unclear what role assignment NP2 adopts. However, reading-time data shows that readers encounter a boggle at the disambiguation point (“began”) suggesting that they had initially assigned incorrect roles (Wilson & Garnsey 2009).

(1) Clause: The two hunters heard the birds began migrating in early October

In contrast, in other types of constructions, grammatical role assignment is delayed. For example, in ORCs like (2b), there is a similar period of ambiguity but reading-time data shows no boggle at the disambiguation point (NP2 of the embedded clause) (Gordon et al. 2001). This suggests that that comprehenders wait until the verb to assign grammatical roles.

Context: The blue bear pushes the horse. The horse pushes the yellow bear.

(2a) SRC: The bear that pushed the horse ate the sandwich. [eater = agent bear]
(2b) ORC: The bear that the horse pushed ate the sandwich. [eater = patient bear]

What accounts for these differences in the timing of grammatical role assignment? One possibility is that role assignment relies on knowing the syntactic structure of the unfolding sentence, and knowledge of the syntactic structure relies on knowledge of the verb. Thus, comprehenders must wait until they encounter the verb before they can assign grammatical roles (SINGLE SOURCE). This can lead to incremental interpretation when the verb occurs early in the sentence but delayed interpretation when the verb occurs late in the sentence. A second possibility is that comprehenders can use information from the verb as well as from noun position (DUAL SOURCE). Because English is an SVO language, NP1 in an English sentence is most likely an agent. However, if NP1 is incorrectly characterized, comprehenders may need to wait for the verb to reassign grammatical roles. This account is consistent with work that shows that information from the noun influences role assignment in relative clauses, but its effects are independent from the time course of structural information from the verb (Lowder & Gordon, 2012). A third possibility is that comprehenders can make use of not only noun position and verb bias to infer grammatical roles, but also any function words that might inform the identity of the construction (ALL SOURCES). This account provides a natural explanation for the rapid disambiguation in garden path sentences because it correctly predicts comprehenders’ tendency to commit to role assignments well before linguistic disambiguation. The current study presented listeners (n=41) with animated events (see “Context” above), followed by sentences like (2a-b). Analyses of eye-movements focused on the proportion of looks to the patient minus agent. During the ambiguous and embedded clause regions, equivalent preferences for the agent emerged in both conditions. Critically, eye-movements appropriately diverged in the main clause region, with increased looks to the agent in the SRC trials and the patient in the ORC trials. These results demonstrate that comprehenders were initially assigning the role of “agent” to NP1, but in constructions that required reanalysis (ORCs, passives) these initial assignments were revised at the verb. This is consistent with the dual source account and suggests that comprehenders use both noun position and verb information to infer grammatical roles, and the availability of this information guides the time course of interpretation.

Table 1. Average agent preference by region and condition.

<table>
<thead>
<tr>
<th></th>
<th>Ambiguous region</th>
<th>Embedded clause</th>
<th>Main clause</th>
<th>Instruction offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC</td>
<td>0.24</td>
<td>0.16</td>
<td>0.26</td>
<td>0.42</td>
</tr>
<tr>
<td>ORC</td>
<td>0.27</td>
<td>0.13</td>
<td>-0.15</td>
<td>-0.28</td>
</tr>
<tr>
<td>p-value</td>
<td>p &gt; 0.50</td>
<td>p &gt; 0.50</td>
<td>p &lt; 0.001</td>
<td>p &lt; 0.001</td>
</tr>
</tbody>
</table>
READERS WITH LESS COGNITIVE CONTROL ARE MORE AFFECTED BY SURPRISING CONTENT
Bruno Nicenboim, Shravan Vasishth & Reinhold Kliegl (University of Potsdam)
bruno.nicenboim@uni-potsdam.de

A series of studies (DeLong et al., 2005; Van Berkum et al., 2005; and others) have shown an effect of gender, phonology, or animacy expectancy at the determiner or adjective prior to a predictable noun. These studies showed facilitation for predictable words via speedups (or reduction of N400 or similar ERP components) before postlexical contextual integration is possible, providing strong evidence in support of a top-down mechanism of predictions.

A mechanism of predictions implies that the parser has to build up potential upcoming continuations, store them, and then evaluate them given the actual continuation (bottom-up evidence). Although it is clear that individual differences should affect this mechanism of predictions, the results are not conclusive (see Otten & Van Berkum, 2009). We show that cognitive control modulates reading times (RT) when evidence mismatches the predicted.

We used a paradigm similar to the one used in Van Berkum et al’s study, taking advantage that determiners in German are always gender marked. As a first step, we conducted a completion task (N=38) with sentences such as “Der Bäcker schob den Teig in ...” (The baker put the dough in ...). We computed a probability score for each determiner and selected 26 sentences with the highest difference between the most likely determiner form (from .61 to 1) and the less likely one (from 0 to .25), creating items such as the following:

(1) a. Der Bäcker schob den Teig in seinen Ofen und wartete. (predictable)
   The baker put the dough in his.masc oven and waited.

   b. Der Bäcker schob den Teig in seine Mikrowelle und wartete. (unpredictable)
   The baker put the dough in his.fem microwave and waited.

We conducted a self-paced reading study (N=81) with sentences such as (1), comparing RTs at the determiner and using as covariates, cognitive control as measured by an adaptation of De Houwer’s (2003) version of the Stroop task, working memory as measured by the operation span task, and reading skills as measured by rapid automatized naming.

Predictions for the critical region (gender marked determiner):
(I) Shorter RTs for the predictable determiner in comparison with the unpredictable one.
(II) RTs for the unpredictable condition should get shorter in comparison with the predictable condition as cognitive control increases, since cognitive control ability is related to the suppression of irrelevant information (Gernsbacher, 1990, 1997; Boudewyn et al. 2012).
(III) RTs for the predictable condition should get shorter in comparison with the unpredictable one as working memory capacity increases, due to the increased ability of high-capacity readers to maintain the predictions (Otten & Van Berkum, 2009).
(IV) RTs for the predictable condition should get shorter in comparison with the unpredictable condition as reading skill increases, since skilled readers are more sensitive to the semantic cues available to them (Pearlmutter & MacDonald, 1995).

Results and conclusions:
* Significant cognitive control × condition interaction (using residual RTs or including word length in the model: t=2.4). As cognitive control decreases, the RTs at the determiner for the unpredictable condition get larger, showing that the conflict between predictions and evidence is modulated by cognitive control.
* Numerical difference between the conditions, which is consistent with the results from the literature. There were no significant interactions between reading skills or working memory and condition, probably because the sentences were too short and simple.

The novel finding in our study is that readers with less cognitive control are more affected by surprising content, i.e. a determiner such as seine (fem) in (1) that does not match the expected (masc) gender according to the predictions for the upcoming noun (masc: Ofen). The longer RTs of low cognitive control readers can be explained by their increased difficulties in overcoming the mismatch between predictions and evidence. The results suggest that cognitive control modulates the ability to suppress irrelevant predictions, in this case the ability to inhibit the predicted gender form when it is not supported by evidence.
CERTAIN PEOPLE DON'T NEED MODIFICATION: THE EFFECT OF INDEFINITENESS ON JAPANESE RELATIVE CLAUSE ASSOCIATION PREFERENCES

Chunhua Bai, Douglas Roland, & Yuki Hirose (The University of Tokyo)
bchjp@phiz.c.u-tokyo.ac.jp

While it seems obvious that the degree to which a noun phrase is likely to be modified plays an important role in resolving attachment ambiguities, it is not always clear which properties of a noun may cause it to attract modification. We argue that in Japanese, bare common nouns inherently attract modification. This means that in the Japanese relative clause attachment ambiguity (which has the form RC + N1-GEN + N2), when N1 and N2 are both bare common nouns (as is the case in most experiments), there is an enhanced bias towards the N1 interpretation, as comprehenders feel the need to modify the bare N1 (N2 does not face such pressure, as it is inherently modified by N1).

Previous results suggest that a temporary local N1 interpretation is established, with comprehenders ultimately settling on an N2 interpretation (Kamide & Mitchell 1997; Miyamoto et al. 2004). Yamada et al. (2014) provide on-line evidence for both an initial N1 interpretation and a subsequent N2 interpretation, even in at least some situations where one of the interpretations seems ill-advised. They use these results to argue for a processing preference towards N2 modification. However, such a preference does not explain the initial N1 interpretation – which we argue is at least partially due to the bare N1 attracting modification.

One reason why bare nouns may attract modification is that in Japanese, noun definiteness is not morphologically marked. Thus, in the absence of a preceding context (as is the case in typical experiments), the comprehender may have difficulty trying to establish the referent of a bare common noun (e.g., woman), and attempt to modify it. To reduce this pressure, we exploited the prenominal indefinite demonstrative aru (some or certain). Aru is typically used by speakers to indicate that the referent of the noun is not in common ground, and that the exact identity of the referent is not relevant to the discourse. As a result, the pressure on the comprehender to find modification for the noun is reduced.

We conducted a self-paced reading study, using aru to signal to the comprehender that the N1 can be interpreted without further information. We used a 2x2 design: N1 type (indefinite (with aru) vs. bare noun) x RC modification type (RC normed to be semantically compatible with both N1 and N2 vs. RC compatible with N1 but not with N2), with 24 items and 38 subjects. The 48 fillers included cases where RC was only compatible with N2 to counterbalance the N1 only experimental items. The reading time for the critical region (N1 + N2) exhibited the predicted interaction (β = -66.78, SE = 25.74, t = -2.59, p < .001). Semantic incompatibility between the RC and N2 was reflected in increased reading times compared to the baseline neutral condition only when the N1 was accompanied by aru. This provides evidence for an early N2 preference when N1 was clearly marked as indefinite, and thus under no pressure to take further modification from the prenominal relative clause.

<table>
<thead>
<tr>
<th>Definiteness</th>
<th>RC modification</th>
<th>RC head noun (critical region)</th>
<th>region3</th>
<th>region4</th>
<th>region5</th>
</tr>
</thead>
<tbody>
<tr>
<td>indefinite</td>
<td>N1 only</td>
<td>titoiya ni natta father became</td>
<td>aru dansei modoru no hahaoya wa certain male model-GEN mother-TOP</td>
<td>totemo kireida beautiful soudesu seem to be</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N1 or N2</td>
<td>byouki ni natta sick became</td>
<td>aru dansei modoru no hahaoya wa certain male model-GEN mother-TOP</td>
<td>totemo kireida beautiful soudesu seem to be</td>
<td></td>
</tr>
<tr>
<td>bare</td>
<td>N1 only</td>
<td>titoiya ni natta father became</td>
<td>dansei modoru no hahaoya wa male model-GEN mother-TOP</td>
<td>totemo kireida beautiful soudesu seem to be</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N1 or N2</td>
<td>titoiya ni natta father became</td>
<td>dansei modoru no hahaoya wa male model-GEN mother-TOP</td>
<td>totemo kireida beautiful soudesu seem to be</td>
<td></td>
</tr>
</tbody>
</table>

*raw reading time (ms)*

- Definiteness: Certain people don’t need modification.
- RC modification: N1 only vs. N1 + N2.
- RC head noun: The mother of a certain male model who became father seems to be very beautiful.
- Region: Raw reading time (ms).
Recent research has revealed strong similarities between the processing of musical structure and linguistic structure. In both language and music, people implicitly acquire a strong expertise in structurally processing sound sequences, based on daily exposure from early infancy. Therefore, recent theories such as the Shared Syntactic Integration Resource Hypothesis (SSIRH, Patel, 2003) suggest that the abstract sequencing processes - using the domain specific rules to parse and structure the sequence - might be shared between domains.

Though previous studies (e.g. Slevc, Rosenberg & Patel, 2009; Perruchet & Poulin-Charronat, 2013) have found interference effects between the two domains, these studies had the limitations of being based solely on linguistic measures of sentences containing structural difficulties. This leaves open the question of whether it is possible to generalise to mutual interference when regarding more direct measures of natural structural processing. Therefore, this study investigated the effects of linguistic unexpectancies on the direct harmonic processing of naturally structured melodies.

Participants heard 8-tone melodies that had 6 regular tone transitions and one harmonic change, which imposed a boundary in the melody. When testing the recognition of a two-tone probe, it will be easier for participants to recognize two tones with regular transition as occurring sequentially, than to recognize the two tones separated by the harmonic boundary (cfr. Tan, Aiello & Bever, 1891). With each tone, a sentence fragment was presented on the screen, so that participants simultaneously read sentences while listening to the melodies. After presentation of these stimuli, the two-tone recognition probe was presented. Sentences were correct (“Tell the assistant that the doctor examined the patient”), or included either a sentence error (“...doctor kitchen the patient”) or a syntactic (“...doctor examined by the woman”) garden path unexpectancy. The sentence fragments containing these irregularities were provided either simultaneously with the harmonic change onset, or at a different point in time.

Recognition of the “harmonic change”-probe was indeed harder than recognition of the non-change probe in all conditions. Only when regarding the data of syntactic unexpectancies (where structural reintegration led to a correct sentence), there was an interaction showing less difference in “probe” type when the syntactic unexpectancy was presented simultaneously ($\beta=0.6718$, $z=2.671$, $p=0.0076$). This provides evidence for the idea that the harmonic integration of the melody (processing the harmonic change) requires integration resources shared with linguistic syntax processing, as suggested by the SSIRH (Patel, 2003).

This study presents online parsing data from children’s subject and object relative clause processing in Greek. Although numerous studies have investigated children’s off-line comprehension of relative clauses in various languages (see e.g. Arnon, 2005; Arosio, Adani & Guasti, 2009; Kidd & Bavin, 2002), very little research has been conducted online. In adult online processing, the difference in processing difficulty between subject relative clauses (SRCs) and object relative clauses (ORCs) appears to be strongly connected to the configurational properties of the language under investigation (Kovács & Vasishth, 2013). Particularly interesting for the present study is the finding that, in free word order languages, online relative clause processing depends strongly on the linear distance between the RC verb and the relative pronoun independently of whether the structure is an SRC or ORC (Levy, Fedorenko & Gibson, 2013).

We present data from an online self-paced listening task in Greek, a free word order language, in which we manipulated the type of relative clause (SRC vs. ORC) and the RC internal word order (canonical vs. scrambled). We recorded online listening times and grammaticality judgments from 24 children (11–12 years old) and 40 adults. Our experimental sentences included RCs introduced by the relative pronoun o opios ‘who’ (nominative case in SRCs, e.g. O majiras-NOM o opios-NOM esprokse ton servitoro-ACC ekapse to fajito, ‘The cook who pushed the waiter burned the food’, accusative case in ORCs, e.g. O majiras-NOM ton opio-ACC esprokse o servitoros-NOM ekapse to fajito, ‘The cook who the waiter pushed burned the food’). Each sentence was presented in a segment-by-segment fashion, and in the end of every sentence participants were asked to judge the grammaticality of the sentence by pressing one of two buttons.

The analysis of participants’ mean listening times revealed no main effect of relative clause, and thus no SRC vs. ORC preference for either children or adults. This reflects findings from two other free word order languages (Levy et al., 2013; Kovács & Vasishth, 2013). A significant two-way interaction between type of RC and word order indicated that local configurations (verb adjacent to relative pronoun) are preferred over non-local ones. The lack of difference between the two groups of participants indicates that - at least for free word order languages such as Greek - children at the age of 11–12 have already developed parsing strategies similar to those of adults (Clahsen & Felser, 2006).

References
Vasishth and Lewis (2006) have proposed that reactivation of upcoming VPs by adjuncts, and/or reactivation of arguments by intervening adjuncts might account for anti-locality effects: facilitation at the verb with increasing distance. An alternative explanation for anti-locality effects is that the expectation of upcoming verb is sharpened as distance increases (Levy, 2008). Two Hindi self-paced reading studies (n=82 in each) show that expectation and reactivation effects interact.

Expt 1 manipulated whether an NP (the subject of a matrix verb) is reactivated, by either making this NP a subject of an embedded nonfinite clause or not:

(1) NP [… nonfinite-verb] … matrix-verb…

Under the reactivation account, when there is a dependency between the NP and the nonfinite verb, the NP is reactivated, leading to faster access of the NP at the matrix verb. This two-level NP-reactivation factor was crossed with a VP-reactivation factor: additional adverbials were either present or absent before the matrix verb that reactivated the matrix verb. This 2x2 factorial design was further crossed with a context factor (prediction context vs no-prediction context) that makes the critical matrix-verb (and the NP subject) either completely predictable or not. This gives us a 2x2x2 design, crossing NP-reactivation, VP-reactivation, and predictability of the matrix verb/NP subject. Results show a significant interaction (t=2.14) between the three factors, such that NP- and VP-reactivation jointly lead to a facilitation at the matrix verb in the no-prediction context condition only; in the prediction condition, no facilitation is seen at the matrix verb. This experiment points to a model of processing where memory activation and expectation are not mutually exclusive factors but closely interact.

Expt 2 replicated the results of Expt 1; here, activation of the critical nonfinite verb was manipulated by placing an adjunct that either modified (i.e., reactivated) the nonfinite verb (NFV-reactivation) or the matrix verb (MV-reactivation):

(2) [NP adjunct … nonfinite-verb] … matrix-verb…

Like in Expt 1, this two-level reactivation factor was crossed with distance (long vs short); long distance conditions had additional adverbials before the nonfinite verb that, by assumption, further reactivate the nonfinite verb in the course of modifying it. This 2x2 factorial design was crossed with a context factor as in Expt 1 to yield a 2x2x2 design; the context ensured that the nonfinite verb was either predictable (prediction context condition) or not predictable (no-prediction context condition). Results show a significant interaction (t=-2.04) between the three factors, such that there is facilitation at the nonfinite verb due to increased distance in the NFV-reactivation, no-prediction conditions; but no facilitation at the nonfinite verb in the prediction conditions. This replicates Expt 1’s findings.

This is, to our knowledge, the first set of studies to show the interaction between activation and expectation strength. Experiments 1 and 2 show that when a weak expectation is raised (i.e., when a verb is predicted but its exact identity is not predictable), reactivation effects are seen, and when a strong expectation is raised (i.e., when the exact identity of a verb is predictable), reactivation effects disappear. We propose that the effect of strong vs weak expectation should lead to relatively high vs low activation of the the NP-subject and the upcoming VP; when activation is high (i.e., in strong expectation conditions), reactivation does not have any effect; when activation is low (i.e., in weak expectation conditions), reactivation shows a facilitatory effect.
PRIMING PROSODY: SPEECH RATE AND BOUNDARY PLACEMENT
Kristen M. Tooley (Texas State University), Agnieszka E. Konopka (Max Planck Institute for Psycholinguistics), and Duane G. Watson (University of Illinois and Beckman Institute)
ptooley@txstate.edu

An ongoing question for theories of spoken language production is whether different aspects of prosody, such as intonational boundaries, speech rate, intonation, and pitch accenting are controlled by the same underlying representations. Contrary to this, recent work suggests that some aspects of prosody, like speech rate, can be primed (Jungers & Hupp, 2009) while others, like intonational boundaries, cannot (Tooley, Konopka, & Watson, 2014). This raises two possibilities. Either these two aspects of prosody are planned at different stages of processing and thus are likely controlled by different levels of representation, or this discrepancy reflects differences in experimental contexts.

To evaluate these possibilities, we used a prime-target paradigm to test whether boundary placement and speech rate of prime sentences can influence the production of new target sentences in one study. We orthogonally manipulated prime sentences (via cross-splicing) to either have an intonational phrase boundary or not (e.g. The monkey that stole the hat // refused to give it back) and to be either 10% faster or 10% slower than the natural speech of the original recordings (by resynthesizing them using Praat). Prime and target sentences always had the same syntactic structure.

64 participants listened to 20 prime sentences and repeated them aloud. One target trial followed each prime, where speakers silently read a novel sentence and repeated it aloud (Boundary Condition, Speech Rate, and Prime-target Order were fully counterbalanced). These trials were embedded in a list of filler trials, split evenly (and randomly) between listen-repeat or read-repeat sequences. Participants’ boundary productions were assessed in two ways: one coder rated whether a boundary was discernible in the critical region or not, and a second coder measured the duration of the word produced before the boundary location through the onset of the first word following the boundary location. Total speaking time of each sentence was also measured.

The analysis of overall speaking times showed significant main effects of Boundary and Speech Rate, and a significant interaction between Speech Rate and Prime-target Order (all ps<.05). Participants tended to speak slower when they heard a boundary, and also spoke faster or slower when primed to do so, but this effect was smaller in the targets than the primes.

The analyses of perceived pauses and word-and-pause durations showed main effects of Boundary and Speech Rate, and interactions between Boundary and Prime-target Order (all ps<.05). Participants were more likely to produce a pause after hearing a slow-rate prime. They also paused at the critical region more often when they were primed to do so: this effect was highly reliable in the prime sentences, but only a weak effect persisted into the targets.

The results show that speech rate but not boundary placement can be primed across sentences, which is consistent with both Jungers and Hupp (2009) and Tooley et al. (2014). Thus rather than being an artifact of differences in experimental contexts, different cognitive mechanisms may underlie the planning and production of speech rate and boundary placement.
PREDICTION GONE WRONG: A MOUSE-TRACKING STUDY IN SENTENCE COMPREHENSION
Franklin Chang (University of Liverpool) & Agnieszka E. Konopka (MPI for Psycholinguistics)
franklin.chang@liverpool.ac.uk

Prediction is thought to enhance sentence comprehension (Pickering & Garrod, 2007), because it is supported by a smart system that is mostly right and therefore likely to speed up processing. An alternative view is that prediction does not always benefit comprehension because, on average, predictions are wrong more often than they are right. On this view, prediction may be better viewed as part of processes that support error-based learning (Chang, Dell, & Bock, 2006).

To compare these accounts, a mouse-tracking study was carried out where 42 participants selected one of four objects on a computer screen based on written instructions like “Click on the car that is orange” (Fig.1). Target displays included a goal object (e.g., \textit{orange car}), a distractor object in a different color (\textit{pink car}), and two other objects, one of which matched the color of the goal object (\textit{orange drawer}, \textit{blue} walnut). Target trials were preceded by 3 types of primes where participants selected the same goal object as in the target trials based on instructions presented with and without modifiers (15 trials per condition). \textit{Neutral} primes had goal objects of a different color than those on target trials (“Click on the car [that is red]”). \textit{Match} primes had the same color goal object in both prime and target displays (“Click on the car [that is orange]”). \textit{Mismatch} primes had goal objects that matched the color of a distractor in the target trial (“Click on the car [that is pink]”).

Reaction times (RTs) for clicking on the correct goal object were compared across conditions over time. If predictions are generated by a smart comprehension system, then Neutral RTs should be intermediate between Match RTs and Mismatch RTs. If Neutral RTs pattern like Mismatch RTs, this would suggest that predictions in this condition can routinely slow down processing in a manner akin to when the system is explicitly misled.

RTs dropped over the course of the experiment ($t=-8.8$), likely reflecting the learning of expectations that supported fast responding to instructions (Fig.2). Importantly, Match RTs were faster than both Neutral and Mismatch RTs ($t=-3.8$), but Neutral and Mismatch RTs did not differ from each other. There was no difference due to presence of the modifier.

Thus in the Neutral condition, participants searched for the goal object of the previous prime trial, even though an object of that color was not present in the target displays. This predictive searching led to a slowing of performance that resembles the slowing on trials where the prime goal object matched a distractor object in target displays (Mismatch condition). Since the Neutral condition best reflects real-world comprehension situations (it requires the least coordination of expectations and the world: e.g., searching for apples, then oranges, then bananas in a grocery store), the fact that it is easy to hinder performance in these conditions suggests that prediction is not necessarily supported by a smart adaptive system but that it supports other functions like learning (Dell & Chang, 2013).
Durational cues are used in language to fulfil several different functions; many languages use the contrast between short (singleton) and long (geminates) consonants to differentiate lexical meaning (e.g. Bengali [pata] ‘leaf’ vs. [pat:a] ‘location’). How these durational aspects are represented and how the representations, in turn, affect the processing of length contrasts has not yet been definitively established. Previous research shows that phoneme quality and quantity are processed independently of each other (cf. Ylinen et al. 2005), which has given support to theories proposing a separate level of representation for non-featural contrasts (e.g. duration and tonal accent). To gain a better understanding of the nature of the representations of durational contrasts, and how these underlying differences may affect processing, we conducted an MMN study in Bengali, a language with a productive medial geminate-singlet contrast.

Previous experiments - lexical decision tasks with fragment form priming and full-word semantic priming - showed an asymmetric pattern of results in both the behavioural and ERP (N400) data. A geminate nonword (‘[ken:a]’) is accepted for a real word singleton ([kena]) and elicits a significant priming effect as well as no enhanced N400 response, but a singleton nonword (‘[g'ena]’) does not result in facilitation of a target related to a geminate word ([g'ena]) and displays a significantly greater N400 amplitude.

A subsequent mismatch negativity (MMN) study investigated whether this asymmetry is also evident in pre-attentive auditory processing. Bengali word/nonword pairs which only differed in the duration of the medial consonant ([g'ena]/[g'ena] and [kena]/[kena]) were presented in a standard oddball paradigm (15% deviants). The results show a latency difference with the singleton nonword ‘[g'ena]’ being significantly slower than the real word geminate [g'ena] while [kena]/[kena] peak at similar latencies. The asymmetry is thus already evident in pre-attentive processing despite the distance of deviance in the stimuli being identical (which is evidenced by the lack of amplitude difference between conditions).

In the case of a nonword, where there is no available lexical representation, the MMN response is slower than when a lexical entry can be accessed. This is in line with previous evidence that the MMN also reflects higher cognitive processes and access of linguistic long-term memory traces (cf. Näätänen et al. 2001) and lends support to the theory that the singleton nonword ‘[g'ena]’ is treated as a nonword while the geminate nonword *[ken:a] elicits the same pattern as the corresponding real word [kena].

These findings provide greater insight into the processing of linguistic duration and how this is mapped onto a representation of length in the mental lexicon. Geminates are represented by a length specification (e.g. mora) on the prosodic level while singletons are not. A geminate mispronunciation subsumes the singleton real-word representation because all other (featural) information is identical. However, when a geminate is mispronounced as a singleton, the length specification necessary to match a geminate representation is lacking and activation fails. Thus, full lexical access is achieved through a mispronunciation only if there is sufficient duration in the acoustic signal to map onto the length specification of the corresponding real-word.

REFERENCES


Anecdotal evidence tells us that people differ in their ability to follow a conversation in noisy environments such as cafeterias or bars. While many studies have compared hearing in noise (HIN) performance between younger and older participants or healthy and clinical populations, researchers have also started to investigate individual differences in healthy individuals contributing to HIN. Because hearing acuity, measured, e.g., by pure-tone hearing threshold, is not a good indicator of HIN, other factors are likely to contribute to individual performance. For example, individuals may differ in their ability to attend to the relevant speaker and ignore irrelevant background babble and a larger working memory capacity (WMC) may enable individuals to devote more attentional resources to the encoding of speech in noise (Akeroyd, 2008; Rönnberg et al., 2013). Recent studies have shown that cognitive variables may indeed contribute to individual differences in HIN (Anderson et al., 2013; Tamati et al., 2013).

The present study sought to find individual differences that contributed to HIN ability in a sample of healthy young adults. Thirty-eight participants completed a HIN test on which they heard 128 sentences (taken from the revised speech in noise test) at two signal-to-noise ratios (+2dB and -3dB; multitalker babble) in predictive and nonpredictive contexts and were asked to type the last word. In addition, they completed two standardized language tests (vocabulary knowledge and verbal reasoning, assumed to reflect verbal ability (VA)), a verbal WM test (two subtests, taken from the NIH toolbox), a test of auditory attention (AA; adapted from Zhang et al., 2012), and a consonant discrimination test (CD; modeled after Garcia Lecumberri & Cooke, 2006). The WM test required participants to repeat items (animals or food) of varying set sizes in size order from smallest to biggest. In the AA test, participants heard two tones and had to decide as quickly and as accurately as possible whether they heard them in the same or different ears while ignoring the frequency of the tones, which was either the same or differed as well. A score for each participant was calculated by subtracting response times on congruent from incongruent trials. In the CD test, participants heard syllables in a VCV cluster presented in multitalker babble and had to identify the consonant by clicking on one of 16 options. The percentage of correct trials was entered as a participant’s score.

Based on previous studies, it was hypothesized that all four variables may contribute to HIN performance. Initial correlation analyses found that the language tests were correlated with WM and CD (rs > .4) and so a structural equation model (partial least squares method) was built to determine the relationship between the various variables. Only VA predicted HIN performance in this model. VA was also a significant predictor of WM and CD (see Fig. 1). To follow up, a binomial mixed-effects regression model on HIN accuracy showed that both noise level and predictive context were significant factors. Furthermore, VA interacted with predictability such that participants with higher VA benefited more from a predictive context. These results suggest that individual differences in HIN in the tested sample could largely be attributed to differences in VA. A larger vocabulary may be associated with more entrenched phonological representations of words in the mental lexicon. A better-entrenched lexicon may make the encoding and rehearsal of words presented in noise less effortful. Alternative hypotheses will be discussed.
INTEGRATION OF "SECONDARY" INFORMATION DURING COMPREHENSION: CO-SPEECH GESTURES AND MOUTH PATTERNS ACCOMPANYING SIGNS

David Vinson, Pamela Perniss, Wing-San Chan, Amy Rice & Gabriella Vigliocco
(University College London). Contact d.vinson@ucl.ac.uk

Numerous studies now highlight multimodal aspects of language comprehension, not only studies in which acoustic/visual aspects of speech influence perception but also those showing how aspects of communication traditionally considered "nonverbal" contribute to comprehension. Here we investigate how information conveyed by a primary linguistic channel is affected by secondary information conveyed visually. In studies of audio-visual speech, crucial evidence about integration comes from studies using incongruent pairings, as in the classic McGurk effect (visual information affects acoustic perception). Similar studies of speech and gesture using incongruent pairings also reveal that gestures are integrated with speech in comprehension (e.g. Kelly, Özyürek & Maris, 2010), but the aim of avoiding visible mouth movements (perhaps along with technical limitations) led to the use of stimuli in which a speaker's gestures, but not head or face, are visible, as in Figure 1. The use of such materials may exaggerate the extent to which gesture is integrated vs. more typical conversational contexts. To address this issue we digitally edited videos to produce materials in which a speaker's face is fully visible, but is combined with a naturally produced gesture that accompanied a different utterance, as illustrated in Figure 2. In addition to considering co-speech gestures we also investigate mouth movements that accompany manual production in British Sign Language (BSL). In BSL, nouns or verbs are frequently produced with lip patterns resembling the production of the corresponding English word albeit without voicing. As the role of these mouth patterns is unclear, we investigate how such "secondary" visual information is integrated into BSL sign comprehension also using digitally altered materials.

In two experiments we presented a picture followed by a video clip and asked participants to indicate whether the word (spoken) or sign (produced by hands) matched it. This provides a strong test of whether the secondary information (gesture accompanying English, or mouth pattern accompanying BSL) is integrated in comprehension, as in both cases the secondary information is not needed to perform the task. Despite this, in both experiments incongruent information increased error rate in "match" conditions (Figures 3 and 4), providing strong evidence for obligatory integration of this "secondary" channel both for speech and signs.

Figure 1: Example of incongruent speech & gesture: Kelly et al. (2010)

Figure 2: Creation of BSL video stimulus. English materials were created similarly (but with audio)

Figure 3: Audio-visual English (n=70 native speakers) Error bars: SE(mean) by subject

Figure 4: BSL (n=24 signers: 16 deaf, 8 hearing) Error bars: SE(mean) by subject
Does switching between linguistic variants (i.e. standard language and a dialect) tax cognitive control mechanisms in similar ways as switching between different languages? Using picture-naming, previous research demonstrated that for bilinguals, switching between languages following unexpected prompts incurs a cost associated with inhibiting one language while activating the other (Costa & Santesteban, 2004). We investigated whether a similar cost is associated with switching between closely related dialect varieties. Despite considerable phonetic and lexical overlap between varieties, do bidialectal speakers inhibit competing phonetic and lexical variants while monitoring whom they can and cannot address in each of the varieties? We hypothesised that if routine use of two similar varieties leads to distinct representations of these varieties then bidialectal speakers have to engage inhibitory control mechanisms, which should incur a switching cost.

To increase cross-linguistic generalisability, we conducted two similar experiments in different regions of Europe: one with bidialectal speakers of Standard Scottish English (SSE) and a regional Scots dialect (n=27), and one with bidialectal speakers of Standard German and a regional German dialect (n=32). Participants named 18 different images depicting cognate (e.g., English: ‘house’ vs. ‘hoose’; German: ‘Eimer’ vs. ‘Emmer’ [bucket]) and non-cognate (e.g. English: ‘children’ vs. ‘bairns’; German: ‘Hose’ vs. ‘Boks’ [trousers]) words. Cognate and non-cognate words were presented in separate blocks of 72 items containing a similar amount of switch and non-switch trials, and a coloured prompt cueing picture naming in one or the other variety.

In both experiments, we found significantly longer RTs for switch trials compared to non-switch trials (see Fig. 1), supporting the idea of inhibitory control associated with producing words from different dialect varieties. We also found faster RTs for cognates indicating facilitation from phonological similarity in line with cascaded models of lexical access (Costa et al., 2000). These findings provide the first evidence for costs associated with switching between closely related varieties. While further research needs to determine to what extent this cost is associated with inhibition of competing lexical entries or competing articulatory settings, the findings show that language production models need to take into account effects of routine use of different dialects of a language. – References:
There is growing evidence that multiword information affects processing (e.g., Arnon & Snider, 2010; Tremblay et al., 2011). Here, we look at the effect of word and multiword frequency on the phonetic duration of words in spontaneous speech to (a) extend previous findings, and (b) ask whether the relation between word and multiword information changes across the frequency continuum. If highly frequent sequences are stored holistically— as argued in many studies of formulaic language (Wray, 2002) then the effect of word frequency should disappear. If, alternatively, increased sequence usage causes a change in the prominence of word and multiword information (Bybee, 2006), we should see reduced effects of word frequency, and increased effects of sequence frequency for high frequency sequences. By examining the relative weight of word and multiword information across the frequency continuum we can look at the interplay between part (word) and whole (multiword) information in processing and representation.

We looked at the effect of word and multiword frequency on the phonetic duration of the middle word of a trigram in spontaneous speech. We extracted trigrams of interest (middle word is noun, no disfluencies, N=8141) from the tagged and time-aligned Buckeye Speech Corpus. We used the corpus to calculate expected duration by summing the expected duration of each sound in the word. We used the combined Buckeye, Switchboard, and Fisher to calculate all part frequencies (unigram, bigram and trigram). We used mixed-effect linear models to conduct our analyses after reducing collinearity. In the first study, we extend previous findings (Arnon & Cohen-Priva, 2013) by showing that trigram frequency affects single word duration (and not only sequence duration), even when controlling for a host of predictability measures [Table 1]. In the second study, we document a change in the prominence of word and multiword information with growing sequence frequency: for highly frequent sequences, the effect of word frequency on duration decreases while the effect of trigram frequency on duration increases [Table 2]. This novel finding shows that repeated usage leads to a growing prominence of multiword information without obliterating the effect of word information. These findings show that multiword frequency effects on processing are robust, found across a range of different trigrams, and are not only predictability effects in disguise. The findings document the growing prominence of multiword information with repeated usage and support expansive views of the lexicon where both words and multiword units are used as building blocks. They undermine the claim that highly frequent sequences are stored holistically. Instead, they show that repeated usage of sequences leads to a growing prominence of the whole, but that part information is still available.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected phonetic duration</td>
<td>0.91</td>
<td>0.012</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word2-Plural</td>
<td>0.028</td>
<td>0.006</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Word2-Freq</td>
<td>-0.035</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pred (word2word1)</td>
<td>0.008</td>
<td>0.001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pred (word2word3)</td>
<td>0.023</td>
<td>0.001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pred (word2word13)</td>
<td>0.015</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Trigram-Freq</td>
<td>-0.011</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Table 1: Trigram effects on duration when controlling for predictability

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected phonetic duration</td>
<td>0.9</td>
<td>0.012</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word2-Plural</td>
<td>0.024</td>
<td>0.007</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Word1-Freq</td>
<td>0.008</td>
<td>0.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Word2-Freq</td>
<td>-0.038</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word3-Freq</td>
<td>0.025</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Bigram1-Freq</td>
<td>-0.013</td>
<td>0.002</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Bigram2-Freq</td>
<td>-0.034</td>
<td>0.002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Trigram-Freq</td>
<td>-0.034</td>
<td>0.005</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word2-Freq X 4th-quartile</td>
<td>0.02</td>
<td>0.005</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Trigram-Freq X 4th-quartile</td>
<td>-0.036</td>
<td>0.011</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Table 2: Changing effect of word and multiword frequency
(MIS)BINDING OF THEMATIC ROLES AND GRAMMATICAL FUNCTIONS IN PASSIVES

Heeju Hwang (Stanford University)
heejuhwang@gmail.com

**Introduction:** Research on sentence comprehension suggests that comprehenders use simple heuristics such as the common correspondence between thematic roles and grammatical functions (e.g. agent=subject) during processing simple transitive sentences. In particular, Ferreira (2003) found that comprehenders often systematically misinterpreted passives, associating the subject with the agent role. Two experiments reported here examine whether speakers also use the common association between thematic roles and grammatical functions as a heuristic during sentence production.

**Experiments:** In both experiments, participants were given two nouns and a verb and were asked to produce active sentences or passive sentences in Korean. Norming study ensured that only one noun is the appropriate agent of the action (e.g. a frog ate a fly versus a fly ate a frog). In **Experiment 1** (n=33), either the agent or the patient was marked with the nominative particle –ika. When the agent is marked with the nominative particle (e.g. frog-ka/ fly/ eat), participants are only allowed to produce an active sentence (e.g. frog- ka fly- lul ate ‘the frog ate the fly’; active condition); when the patient is marked with the nominative particle (e.g. frog/fly- ka/ eat), participants have to produce a passive sentence (e.g. fly- ka frog- hantey was eaten ‘the fly was eaten by the frog’; passive condition). If speakers use the common association between thematic roles and grammatical functions (i.e. agent=subject) as a heuristic during functional processing, they should have difficulty forming the atypical bindings inconsistent with the heuristic (e.g. patient=subject in the passive condition). As a result, participants are expected to commit greater binding errors in the passive condition (e.g. binding the agent with the nominative particle even though the patient is bound to the nominative particle in a given trial). However, if the production system binds thematic roles and grammatical functions based on algorithmic analysis alone, the passive condition should not result in greater binding errors than the active condition. We found that speakers indeed had more difficulty in the passive condition (28% vs. 6%). When the agent entity was marked with the nominative particle (active condition), participants had little trouble linking thematic roles and grammatical functions (6%). However, when the patient entity was marked with the nominative particle (passive condition), participants were likely to bound the agent entity to the nominative particle (22%) or assume the nominative- marked patient entity to be the agent (6%). This suggests that speakers often bound thematic roles and grammatical functions using the common association between the two (i.e. agent=subject) rather than algorithmically processing the patient–subject mapping in production. Importantly, the greater error rates of binding the agent to the subject suggest that: (i) the production system is more likely to assign the agent to the subject in contrast to the comprehension system assigning the subject to the agent (Ferreira 2003) and (ii) the errors in the passive condition cannot be simply attributed to (mis)comprehension processes. In order to see whether the reliability of a structural cue modulates speakers’ reliance on heuristic processing (e.g. MacWhinney et al. 1984), **Experiment 2** (n=20) marked the agent with the dative particle –hantey (passive condition) and the patient with the accusative particle –(lj)ul (active condition). Unlike the nominative particle, the accusative and the dative particle are a reliable cue to the structure to be produced and the thematic role of the noun - because the accusative particle only occurs in the active structure and the dative particle in the passive structure. Quite strikingly, we found that participants did not commit any binding errors in either active or passive condition. This suggests that speakers’ reliance on heuristics is modulated by the reliability of a particle as a cue to a syntactic structure and the thematic role of a noun.

**Conclusion:** The results suggest that in order to better capture the processes underlying sentence production, theories of sentence production as well as comprehension should take account of heuristic processing and its interaction with other factors influencing production such as cue reliability. (*All statistics are significant at equal to or less than .05 level.*)
IS EXPLICIT MEMORY FOR PRIME SENTENCE STRUCTURE NECESSARY FOR LEXICALLY BASED SYNTACTIC PRIMING?

Sarah Bernolet (Ghent University), Paul Boon (Ghent University Hospital), Miet De Letter (Ghent University Hospital), & Robert Hartsuiker (Ghent University).
sarah.bernolet@ugent.be

Syntactic priming is enhanced when prime and target sentences use the same main verb. It has been suggested that this lexical boost to syntactic priming might be an effect of explicit memory: When the same verb has to be used in consecutive prime and target sentences, the repeated head acts a retrieval cue that gives access to the explicit memory of the prime sentence’s surface structure (Chang, Dell & Bock, 2006). In order to investigate this hypothesis, we compared the production of dative sentences by patients with anterograde amnesia and 2 groups of control speakers (one matched for age and one younger student group). Participants had to describe pictures of dative actions. Before each target picture, they heard and repeated a prime sentence (DO-dative, PO-dative or transitive baseline) and one unrelated filler sentence. After each target picture, participants’ memory of the prime sentence was assessed by means of a probe recognition task.

The patients showed severely impaired memory for the syntactic structure of the prime sentence: When the probe was only syntactically different from the prime sentence, the patients correctly rejected the probes in only 36% of the cases. The control groups had a better memory for syntactic structure, but did not perform above chance level. Nevertheless, like in the study by Ferreira, Bock, Cohen & Wilson (2008), all groups showed significant syntactic priming (patients and age-matched controls 15% priming; students 12% priming). The patient group showed no trace of a lexical boost (0%). The boosts observed in the control groups (3% for the controls matched on age; 4% for the students) were not significant and much smaller than the boost we observed in our pilot study with a student group (20%). Only the patient group showed a significant effect of target verb bias in the baseline condition: The odds for a PO dative response decreased together with the target verbs’ bias towards this response (left panel Figure). In the two other groups this effect was less pronounced and not significant. In sum, the only difference in the production data between the patient group and the control groups was that, in the baseline condition, the patient group showed a stronger effect of target verb bias, maybe because the patients’ responses were less influenced by explicit memory processes.

Are Thematic Roles Linked to Grammatical Function and Linear Order in the Same Way? Evidence From Structural Priming Effects in Japanese Sentence Production

Ying Deng, Douglas Roland & Yuki Hirose (The University of Tokyo)
dengying@phiz.c.u-tokyo.ac.jp

Structural priming effects have been observed in head-initial language as well as head-final languages, such as Japanese, in both comprehension and production studies (e.g., Yamashita et al, 2005). A substantial number of studies report a voice priming effect (i.e., active/passive), and in free word order languages (i.e., Japanese) a word order priming effect is also observed (e.g., Bock, 1982; Tanaka, 2007). Recent empirical evidence suggested that in language production process, the voice priming effect is assumed to occur when the thematic role is mapping to grammatical function, on the other hand the word order priming is assumed to occur when the thematic role is mapping to linear word order (Cai et al., 2012). Since in most previous research the voice priming effect and the word priming effect are observed in separate experiments, less is known about whether the thematic roles are mapping to grammatical function (the priming effect of voice) and linear word order (the priming effect of linear word order) in the same way and whether the magnitude of priming effect is different between voice priming and linear word order priming.

The current study used Japanese, in which the scrambled active (OSV-active) and the canonical passive (SOV-passive) pairs share the same “patient-agent” thematic role order, and therefore allow us to test whether the magnitude of priming effect is different between voice priming and linear word order priming while eliminating the effect of the thematic role order. We conducted two oral picture description experiments. First, Experiment 1 examined whether the priming effect of linear word order reported in the literature can be replicated. Then Experiment 2 compared the magnitude of voice priming effect and word order priming effect within a same experiment.

In Experiment 1, participants were asked to read a prime sentence aloud, and then silently read a verb appearing in its citation form on the screen. After the target picture appeared, participants were asked to describe the picture using the verb given earlier. The 18 items consisted of two types of prime sentences: active sentences with canonical (SOV-active) or scramble word order (OSV-active), and target pictures that involve two human entities for both agent and patient. The results showed that more OSV-active sentences were produced after OSV-active primes (0.14) than SOV-active primes (0.03), suggesting a linear word order priming effect.

In Experiment 2, the primes were OSV-active or SOV-passive sentences. First, we found a priming effect of linear word order, with more OSV-active sentences being produced after OSV-active primes (0.17) than SOV-active primes (0.03), and a voice priming effect, with more SOV-passive sentences being produced after SOV-passive primes (0.42) than OSV-active primes (0.02). Second, to test whether the magnitude of priming effects are different between voice priming and linear word order priming, we combined the data from the both experiments and carried out an analysis with voice priming and linear order priming as predictors. Again we found both significant main effect for voice and linear order priming. More importantly, the results showed that the effect of voice priming (β= 1.75) is greater than word order priming (β= 1.17).

Taken together our results suggest that mapping of the thematic role to grammatical function (voice priming) might have a stronger influence on production process than mapping thematic role to linear word order (linear word order priming) when other factors (thematic role order) is controlled.

**SOV-active prime:** hiroin-ga akuyaku-o yattsuke-ta (The heroine beat the villain.)
hiroin-NOM villain-ACC beat-PAST

**OSV-active prime:** akuyaku-o hiroin-ga yattsuke-ta (The heroine beat the villain.)
villain-ACC hiroin-NOM beat-PAST

**Passive prime:** akuyaku-ga hiroin-ni yattsuke-rare-ta (The villain was beaten by the heroine.)
villain-NOM heroine-OBL beat-PAST-PASSIVE
AN ELECTROPHYSIOLOGICAL ANALYSIS OF THE TIME COURSE OF PHONOLOGICAL AND ORTHOGRAPHIC ENCODING IN WRITTEN WORD PRODUCTION

Qingqing Qu (Chinese Academy of Sciences, Beijing, China) & Markus F. Damian (University of Bristol, United Kingdom)
quqq@psych.ac.cn

In orthographic output tasks such as handwriting and spelling, it is widely accepted that orthographic codes can be accessed directly from semantics (Rapp et al., 1997). At the same time, many recent studies (e.g., Qu et al., 2011) have demonstrated a phonological influence on orthographic encoding; hence, individuals appear to generate orthographic codes based on a direct route from semantics, as well as via indirect input from phonology. The present study used event-related potentials (ERPs) to examine the relative time course of spelling and sound activation in handwriting. We adopted Chinese as the target language because in non-alphabetic orthographic systems of this type, spelling and sound are largely dissociated. Chinese participants were presented with coloured line drawings of simple objects, and wrote colour and picture names as adjective-noun phrases. Colour and picture names were chosen to be either phonologically related (they shared a rhyme), orthographically related (they shared an orthographic radical), or completely unrelated.

EEG analyses demonstrated that phonological overlap modulated the P2 component, hypothesised to reflect early extraction of phonological information (e.g., Carreiras, Vergara, & Barber, 2005), with less positive ERP amplitudes in the phonologically related compared to the unrelated condition. Phonological effects emerged in the 200-500 ms time window (onset latency: 202 ms) across a large set of posterior electrode sites. Orthographic overlap modulated the N250 component which has been suggested to be sensitive to orthographic overlap (e.g., Dunabeitia, Molinaro, Laka et al., 2009) with less negative ERP amplitudes in the orthographically related compared to the unrelated condition. Orthographic effects were found in a 300-400 ms time interval (onset latency: 310 ms) broadly across left-middle regions. Overall, our findings indicate that in written word production, phonological codes are accessed approximately 100 ms earlier than orthographic codes. The current results support the notion that access to phonology from meaning is efficient and speedy and has “processing priority”. By contrast, access to orthographic codes appears to be slower, possibly due to the fact that this link has undergone less practice across the lifetime than retrieval of corresponding spoken codes. Hence both routes contribute to orthographic encoding (as evidenced by the independent behavioural and ERP priming effects obtained for orthographic and phonological overlap in handwriting), but with distinct time courses.

References


Every time bilinguals speak, they must decide which language to produce. Despite the ever-present availability of two lexicons, they are remarkably good at preventing accidental language intrusions (e.g., Gollan, Sandoval, & Salmon, 2011; Poulisse & Bongaerts, 1994). What cognitive mechanisms do bilinguals use to guard against such mistakes? One oft-cited proposal is that bilinguals use language-wide inhibition to reduce the activation of all representations in the non-target language, thereby avoiding cross-language interference (e.g., Green, 1986, 1998; Meuter & Allport, 1989), and that such inhibition must be particularly strong when speaking the non-dominant language. This asymmetric inhibition leads to what is known as the switch cost asymmetry, in which the cost of switching into the dominant language is greater than the cost of switching into the non-dominant language because inhibition takes time to remove (e.g., Allport, Styles, & Hsieh, 1994; Meuter & Allport, 1999; but see Koch, Gade, Schuch, & Philipp, 2010).

If the switch cost asymmetry reflects language-wide inhibition, it should affect all words in the dominant language equally. The present study tested this hypothesis by asking if the asymmetry is equally sized for low- and high-frequency words. Sixty-eight Spanish-English bilinguals switched between naming pictures in their two languages. On every trial, an American flag or a Mexican flag indicated whether the picture should be named in English or Spanish. Bilinguals named 32 low-frequency pictures (mean natural log frequency = 1.98, SD = 0.83) and 32 high-frequency pictures (mean = 4.33, SD = 0.86) once each.

Picture naming latencies for each condition are shown in the Figure separately for low-frequency (LF) and high-frequency (HF) picture names. Results showed a significant switch cost asymmetry, but – crucially – only for low-frequency targets (represented by the difference in slopes for the dashed lines), $\beta = 181.3$, $SE = 63.0$, $\chi^2(1) = 7.70$, $p = .0055$. In stark contrast, switch costs were symmetrical for high-frequency targets (represented by the similar slopes for the solid lines), $\beta = 18.8$, $SE = 37.5$, $\chi^2(1) = 0.24$, $p = .6243$. These patterns for low- and high-frequency targets were significantly different from each other, $\beta = 54.5$, $SE = 22.3$, $\chi^2(1) = 5.74$, $p = .0166$.

As noted above, the modulation of the switch cost asymmetry by word frequency is unexpected under the assumption that inhibition is applied to the language as a whole. Green (1998) also proposed reactive inhibition, which is adjusted proportionally on an item-by-item basis depending on the magnitude of competition. However, this would have predicted a greater switch cost asymmetry for high-frequency than for low-frequency targets (the opposite of the observed pattern). Our results require revisions to notions of how inhibition is used in bilingual language control.
THE INFLUENCE OF SCRIPT KNOWLEDGE ON EXPECTATIONS:
EVIDENCE FROM ERPs
Francesca Delogu (Saarland University), Heiner Drenhaus (Saarland University) & Matthew Crocker (Saarland University)
delogu@coli.uni-saarland.de

When reading a text describing a common activity like going to a restaurant or attending a party, comprehenders build a model of the event described that includes prior knowledge of the participants, locations, and sequences of actions that typically occur within the event. Previous work has demonstrated that such script knowledge, or schemas (Rumelhart, 1980) guides the processing of incoming information by facilitating predictions as to what will occur next (e.g., Zacks et al., 2009; Speer & Zacks, 2005). When predictions are confirmed, the current event model remains active. When predictions fail, and a surprising event is encountered, the current model is updated and an event boundary is perceived.

In the present ERP study we investigated whether the explicit granularity with which such common events are described influences comprehenders’ expectations about what activity could be mentioned next. Participants (N=20) read 120 brief stories such as in the example below. After a short introductory sentence, the context sentence described a common event (e.g., taking a shower) by mentioning either one typical action in the event, like washing the hair (Short context a), or a sequence of actions typically involved in the event, like washing feet, arms and hair (Long context b). The final sentence contained a target word referring to an activity that could either be expected to occur next (shaving) or be initiating a new, unrelated event (hoovering).

We hypothesized new events to elicit larger N400s than more expected events. Assuming that more fine-grained, structured representations of common events elicit higher expectations for a related continuation, we also expected a stronger N400 effect on hoovering following a long context than a short context.

Example of the materials
Intro: Jan ist vom Laufen ganz verschwitzt. (Jan is soaked with sweat because of running).
Short context: Er geht in die Dusche, wo er sich die Haare wäscht. (He goes in the shower, where he washes his hair).
Long context: Er geht in die Dusche, wo er sich erst die Füße, dann die Arme und dann die Haare wascht. (He goes in the shower, where first he washes his feet, then his arms and then his hair.)
Target: Dann begann er mit dem Rasieren/ Saugen, wofür er zehn Minuten brauchte. (Then he began the shaving / hoovering, for which he needed ten minutes).

In the N400 time-window, the results showed larger N400s for hoovering than shaving, but no interaction with context length. Between 600 and 900msec, however, there was an effect of context, with long contexts eliciting a larger frontal positivity than short contexts. Crucially, the effect was much stronger in the new event conditions (hoovering). We interpret the P600 effect as reflecting model updating processes (see Brouwer et al., 2012). Suppressing and updating the current event model when an event boundary is perceived is harder when a more fine-grained, structured representation of the event is held in memory.

References
RETRIEVAL INTERFERENCE DURING COMPREHENSION OF GRAMMATICAL SUBJECT-VERB AGREEMENT: EVIDENCE FROM ERPS

Mante S. Nieuwland (U. of Edinburgh), Ian Cunnings (U. of Reading), Andrea E. Martin (U. of Edinburgh), & Patrick Sturt (U. of Edinburgh)
m.nieuwland@ed.ac.uk

Research on subject-verb agreement during comprehension suggests a ‘grammaticality asymmetry’ in similarity-based retrieval interference. Whereas processing costs incurred by ungrammatical subject-verb agreement are reduced in the presence of a grammatically illicit attractor noun that matches the verb in number, attractor nouns have not been found to affect the processing of grammatical sentences [1]. However, most existing studies have only included singular verbs in the grammatical conditions, and the lack of retrieval interference in such cases could be a result of the fact that singular is an unmarked feature [2]. In the current study, we tested for similarity-based interference for both singular and plural verbs in fully grammatical sentences. If plural is a marked feature, we expect to find evidence of retrieval interference for plural verbs but not singular verbs when multiple items in memory match the number of the verb. We predicted that retrieval interference would elicit a P600 effect [3-4], the effect commonly associated with syntactic processing difficulties.

Methods: Participants read 120 grammatical sentences (30 per condition) belonging to a 2(subject noun: plural, singular) x 2(attractor noun: plural, singular) factorial design in which the critical verb (have/had/were/was) always agreed in number with the subject noun. Sentences were mixed with 280 fillers and presented word by word (300 ms duration, 200 ms blank). Intermittent yes/no comprehension questions were answered with 92% accuracy. EEG data was recorded from sixty-four channels and segmented into epochs from 200 ms before to 1000 ms verb onset. Data was baselined to 0-200 ms post-stimulus to eliminate spurious effects from pre-critical word differences (see also [3-4]).

Results: Using average amplitude per condition across 16 centrally distributed EEG electrodes, repeated measures ANOVAs in the 500-700 ms time window showed an effect of attractor that was reliably different for plural and singular verbs ($F(1,35)=4.8, p<.05$), with a robust P600 effect elicited by plural verbs (PP minus PS voltage difference, $M=+.64, F(1,35)=5.7, p<.05$) but none for singular verbs (SS minus SP, $M=-.22, F(1,35)=.44, ns$).

Conclusions: The observed P600 effect for grammatically correct, plural verbs in context of a plural attractor noun suggests that retrieval interference arises as a by-product of grammatical processing, and constitutes evidence against a grammaticality asymmetry in interference effects.

References
“WHAT IS CONCRETE AND ABSTRACT AND READ ALL OVER!?”: A NEURAL 
DECODING STUDY OF TYPE COERCION IN INHERENT POLYSEMY 

Yuan Tao (University of Trento), Andrew J Anderson (University of Rochester), Massimo Poesio (University of Trento, University of Essex) 
yuan.tao@unitn.it

Coercion has been argued to be an important mechanism of resolving polysemy and meaning mismatch in natural language, and this phenomenon provides us crucial evidence both about the lexicon and compositionality in language. Previous psycholinguistic studies focus on two kinds of coercion: aspectual coercion and complement coercion, and the results generally show that processing coercion incurs an extra cognitive load (e.g. Frisson & McElree, 2008; Kuperberg et al., 2010; Paczynski et al., 2014), implying there is an on-line construction of meaning arising from the verb-noun composition. Studies using MEG have also suggested that brain regions such as the ventromedial prefrontal cortex and the left anterior temporal lobe play an essential role for semantic composition (Pylkkänen & McElree 2007; Bemis & Pylkkänen, 2013).

In this study we used functional Magnetic Resonance Imaging (fMRI) to examine the neural correlates of coercion with a special class of polysemous words, so-called dot-objects -- words like book that have two or more senses systematically related by logical polysemy (Pustejovsky, 1993, 2006, 2011). All different senses of a dot-object are generally accessible, but one sense may be selected in a context by means of a type coercion operation. Such words are of particular interest in light of recent investigations of concept representation in the brain using fMRI, that have identified a general neural distinction between concrete and abstract semantic knowledge (e.g., Binder et al., 2005; Wang et al., 2012). Five words were selected, representative of the archetypal kind of dot-object, “physical object • information” whose meaning can alternate between concrete and abstract in a systematic manner: book, diary, catalogue, magazine, sketch. We compared the neural representation of coerced concrete and abstract senses which were rendered by thirty adjective phrases (e.g. scientific book, worn book). A simple contrast was also constructed in which one abstract and one concrete category were chosen to approximate the partial senses of the selected dot-objects as a comparison (fig.1). The FURNITURE category was chosen because furniture items are often found in the same context as books (e.g. books are often on shelves/tables etc). All stimuli match in number of phonemes and letters, and were normed beforehand to match familiarity and differ in concreteness.

Dot-object theory predicts that the two contrasts should engage similar neural mechanisms. We test this hypothesis using multivariate pattern analysis to distinguish the neural activity patterns of the two categories of each contrast. The simple contrast showed a left-hemisphere dominance, and was most distinguishable in the bilateral precuneus, which is among the regions most consistently activated by semantic processing and typically shows a concrete>abstract effect. (e.g. Binder et al., 2009) On the other hand the concrete and abstract coercions could be distinguished in regions across the bilateral posterior inferior temporal gyrius, which is adjacent to the visual cortex and has been associated with object recognition and representation of visual features (e.g. Martin & Chao, 2001). No region supported classification of both types of contrast as dot-object theory would have expected. These results suggest that the neural representation of the coerced dot-objects is more complex than simply representing the partial sense, highlighting the importance of contextual influence and semantic composition in future studies of concept representation.

Fig.1: Experiment contrasts. During the experiment participants were instructed to attentively think about the target concepts in order to make a semantic relatedness a judgement.
HEMISPHERIC SENSITIVITY TO FREQUENCY AND PREDICTION: EVIDENCE FROM EARLY AND LATE ERPS

Yoana Vergilova (University of Saarland), Heiner Drenhaus (University of Saarland), Matthew Crocker (University of Saarland), vergil@coli.uni-saarland.de

We present an ERP investigation of the contribution of left and right hemispheres in predictive language processing, and its time-course. Dambacher et al., (2009) find early top-down effects of predictability, suggesting that context integration occurs rapidly – in the same timeframe as bottom-up lexical factors such as word frequency or length. In contrast, studies that have examined hemispheric differences have focused on later time windows (e.g. the N400: Wlotko and Federmeier, 2007, 2013) providing evidence suggesting the left hemisphere (LH) emphasises top-down predictive processing whereas the right (RH) is more bottom-up and integrative. In this work we modify Dambacher et al’s (2009) experiment applying the divided visual field paradigm to investigate both the possible interplay of frequency and predictability in the early lexical access timeframe (<250 ms after stimulus onset) as well as their later influences on the N400 in the two brain hemispheres.

Frequency of target words (high/low), predictability of target sentences (high/low) and visual field of presentation of the target words (LVF/RVF) were manipulated examining modulations in early (before 100ms) and later (250-450 ms) ERPs. 19 right-handed native speakers of German saw 144 tandems of context (high/low predictability) + target (same for both high/low frequency words) sentences twice (target words presented once to the LVF/RH and once to the RVF/LH) in pseudorandomised order:

The man on the picture fiddled around with models of Columbus’ fleet. (high predictability)
The man on the picture wore a golden crown and sat on a stately throne. (low predictability)
In his right hand he held a ship (high frequency target) of considerable length.

Early effects (0-100ms after stimulus onset): Based on analysis using linear mixed models (Bates et al, 2008), we report results associated with |λ|>2. We found a frontally distributed positivity for low frequency compared to high frequency words, across both hemispheres, which was in contrast with Dambacher et al. (2009) who reported only an early predictability effect. Our data showed an interaction between frequency and predictability, which revealed that the predictability effect was driven by high frequency. In other words, in our study frequency modulated predictability in the first 100ms. This lends support to the view that top-down and bottom-up factors interact rapidly after stimulus onset (Pulvermüller et al., 2009) with no hemispheric advantage.

Later effect (250-450ms after stimulus onset): The later interval revealed two two-way interactions between predictability and VF and predictability and frequency. The three-way interaction (frequency X predictability X VF) is consistent with the view that modulation of predictability by frequency is mostly present in the LVF(RH), where only high frequency target words exhibited a centro-parietal negativity in high predictability contexts compared to low predictability ones. The low frequency condition did not present this pronounced modulation. Whereas in the RVF(LH) we observed a consistent N400 predictability effect across both frequency conditions.

In conclusion, our study highlights the interplay between bottom-up (frequency) and top-down (predictability) factors at a very early timeframe. We extend Dambacher et al. (2009), who did not report a significant frequency X predictability interaction with central presentation of the same stimuli. Additionally, hemispheres appear to be similarly involved the initial lexical access of words (<100ms). With respect to the N400 time window, our findings are consistent with Federmeier et al. (2013) as the negativity connected to predictability occurs in both hemispheres. Moreover, our results provide additional evidence that the RH is continuously sensitive to frequency, allowing for a possible insight into the nature of top-down/bottom-up hemispheric processing.
OUT OF USE, OUT OF MIND? FIRST LANGUAGE PROCESSING IN A DUTCH 
EMIGRANT POPULATION IN ANGLOPHONE NORTH AMERICA 
Bregtje J. Seton (University of Groningen), Susanne M. Brouwer (University of Utrecht), 
Laurie A. Stowe (University of Groningen) & Monika S. Schmid (University of Essex) 
b.j.seton@rug.nl 

Background: Research in bilingual language processing often focuses on the second 
language of a bilingual population. The first language (L1) of these bilinguals is less well 
studied. The current study addresses the question of whether first language processing 
changes due to emigration to another country, and when the second language has become 
the dominant one, leading to language attrition. Previous research has focused mostly on 
language production and has used only behavioral methods. L1 attriters have for example 
been shown to have difficulties with lexical access (e.g. Schmid 2002), and to make more 
mistakes and have longer reaction times in picture naming tasks (Hulsen et al. 2002). In the 
present study we used an Event-Related Potential (ERP) paradigm to investigate the 
language processing of a group of Dutch attriters, who used English as their dominant 
language. 

Methods: A group of Dutch attriters (N=50, age of emigration: 5 - 48 years old, length of 
residence: 5 - 56 years) and a control group of Dutch native speakers in the Netherlands 
(N=28) took part in an auditory ERP experiment, in which they were presented with 
sentences that contained violations in grammatical gender (e.g., ”de/het huis, the house” or 
in non-finite verb forms (e.g., ”lopen/gelopen, to walk/has walked”). 

Results: The Dutch native speakers showed the typical P600 response to both types of 
violations. In a preliminary analysis of the attriters, the results show that processing the 
Dutch language is still very native-like in most of the attriters who arrived in Canada after 
age seventeen, although all of the participants had become dominant in English and did not 
speak much Dutch anymore. In a small subgroup of attriter participants who arrived between 
the ages of six and twelve, the lack of a systematic P600 effect in the grammatical gender 
condition suggested that their processing was less native-like. This could be due to 
incomplete acquisition of Dutch and a lack of entrenchment of the language. 

Discussion: A shift in language dominance from the first to the second language does not 
apparently mean that language processing of the first language becomes less native-like. 
However, there is large individual variation between participants in our data. The data set is 
being analyzed with respect to a large amount of background variables on language use and 
attitudes which have been shown to affect attrition in earlier studies. With this we hope to 
tease apart the effect of these different variables on the data. 

References 

language shift and language processing in three generations of Dutch migrants in New 

Schmid, M. S. (2002). First Language Attrition, Use and Maintenance: The case of German 
Jews in Anglophone countries. Amsterdam: John Benjamins.
PROFICIENT BILINGUALS’ LEXICAL ACCESS IS AFFECTED BY INTERLOCUTOR CONTEXT

Monika Molnar (Basque Center on Cognition, Brain, and Language; BCBL), Clara Martín (BCBL), Antonio Ibañez (University of Jaen), & Manuel Carreiras (BCBL; Universidad del Pais Vasco; Ikerbasque)
m.molnar@bcbl.eu

In bilingual communities, individuals often use only one of their languages, or they use both of them at the same time (e.g., code mixing). Also, bilinguals seem to adjust to the linguistic background of different interlocutors with ease. In two experiments, we investigated whether bilingual language processing is supported by contextual cues (e.g., interlocutor identity) to achieve such efficiency. In Experiment 1, using an audio-visual task proficient and non-proficient Basque-Spanish bilinguals were familiarized with six novel interlocutors who spoke either Spanish or Basque, or both languages. Subsequently, participants completed an audio-visual lexical decision task, in which the interlocutors produced test items in Spanish or Basque. Proficient bilinguals’ (but not non-proficient bilinguals’) speed of processing increased when the language that interlocutors spoke during familiarization matched the language they spoke at test, relative to test trials when the interlocutors swapped languages, suggesting that proficient bilinguals benefit from an association between language and interlocutor during language comprehension.

In Experiment 2, using the paradigm described above, we assessed whether bilinguals are able to predict the context-appropriate language based on visual interlocutor-identity cues. We measured bilinguals’ brain responses (using event-related potentials; ERPs) between the onset of the visual presentation of the interlocutor and the onset of speech interlocutors produced. The ERPs revealed that around 150 ms after video onset (prior to speech onset), bilinguals exhibit different brain responses depending on whether the trial is presented by a monolingual or a bilingual interlocutor (see Figure). Overall results suggest that bilinguals are able to activate a language (mode) for further speech comprehension relying on contextual cues provided by interlocutor identity.

Figure: ERP results.

Time zero (black vertical line) indicates the presentation of the video of the speaker. Time 300 ms (grey vertical line) indicates the averaged onset of speech.

Black line depicts ERPs measured for Basque interlocutors (who spoke only Basque during familiarization); Dotted black line depicts ERPs measured in response to Spanish interlocutors (who spoke only in Spanish during familiarization); Red line depicts ERPs measured for bilingual identity (who spoke in both Basque and Spanish during familiarization).

Linear derivation of FC1, FC2, C3, Cz and C4 is presented. Negativity is plotted up.
Recent work in psycholinguistics has attempted to isolate lexical complexity as a factor in the processing time for verbs (McKoon and Macfarland 2002, Gennari and Poeppel 2003, Stockall et al 2008, Thompson 2002, McKoon and Macfarland 2002). One common assumption is that alternating verbs like *melt* are underlyingly transitive (Levin and Rappaport Hovav 1995, Chierchia 2004). However, previous work did not directly and minimally compare alternating verbs like *melt/break* with corresponding baseline transitive and intransitive behaviour in the same experiment. Our own experiment is designed to do just this. It also departs from previous work in that we do not presuppose verb classes from syntactic theory (such as internally vs. externally cause verbs) but pose our question directly in terms of the notion of *alternation vs. rigidity in frame compatibility*. Thus, we ask, in an experimental task requiring the processing of verbal ‘subcategorization’ information: do alternating verbs pattern with transitives, with intransitives, or with neither?

**The Experiment:** The experiment consisted of a ‘go/no go’ test, for testing integrating verbs of different types into different syntactic frames (Mauner et al 1995). Subjects were instructed that they would be reading a list of sentences, and told that some of them were possible sentences of English and some not. The task was word by word self paced reading, with an abort button as soon as a sentence was judged not possible. Verbs of four classes were included: strictly transitive verbs (t), strictly intransitive verbs (i), verbs of unstable valency due to causative-inchoative alternation (*break/melt* type verbs) (a), and verbs of unstable valency due to object drop (aa), with 9 verbs in each class. The four classes (and the individual verbs) were matched for frequency (log frequency from Baayen et al.) and word length. The a and aa groups were also matched for transitivity percentage using the BNC. The novelty of the test paradigm was that the crucial test items consisted of a clausal complement frame, where all the test items were ungrammatical. In each such sentence, the verb was followed directly by the complementizer *that*, followed by a name and a finite verb, as in (1a) *The doctor harassed that Kathy was in despair* and (1b) *The doctor speculated that Kathy was in despair*. Thus, the only difference between the ungrammatical sentences in question was the verb class of the verb. Specifically, we checked (i) whether the time to abort (judgement of ungrammaticality) was significantly correlated with verb class and (ii) whether the alternating verbs patterned with transitives, intransitives, or neither, in their behaviour. Three test sets were constructed, and each informant only did one test set. No test used the same actual verb more than once. Three verbs from each verb class (12 in total, all ungrammatical) were included in each test set in the CP test frame, plus 12 grammatical sentences with regular clause-taking verbs for balance. The noun phrases in subject position were chosen from a set of eight animate common nouns, and their relatedness scores to the different verbs tested in a separate judgement task on different speakers.

**Results and Discussion:** Seventy English speaking undergraduates were tested. Our dependent variable was the ‘number-of-words-to-reject’. We used R to perform a linear mixed effects analysis with verb class as the fixed effect. As random effects, we had intercepts for subject and item, and by-subject random slope for the effect of verb class.

* We found that alternating verbs like *melt* (class a) were significantly quicker to reject than transitive verbs, and were non-distinct from the intransitive group. Alternating verbs like *knit* (class aa) took significantly longer than intransitive verbs, and were non-distinct from the transitive group. We thus found no evidence for ‘alternation’ per se being a significant factor affecting processing complexity. We interpret the increased judgement times for transitive verbs as being due to the possible interpretation of the complementizer as a demonstrative.

* What was striking was that *melt*-type verbs patterned like *intransitives* with respect to this effect, and this was so even though the ‘animacy’ of the matrix subject in all cases should have favoured a transitive interpretation of the flexible frame verbs. Verbs like *knit* on the other hand, did pattern with the transitives, showing that the fact of alternation and the proportion of transitive surface alternants in the wild was not the determining factor.
SPATIAL ATTENTION AND IMPLIED SPEED IN SENTENCE COMPREHENSION: PATH AND ADVERB MANIPULATIONS
Shane Lindsay (Dundee), Christoph Scheepers (Glasgow), Yuki Kamide (Dundee)
shane.lindsay@gmail.com, y.kamide@dundee.ac.uk

One account of language comprehension suggests that we invoke perceptual-motor simulations of the events described in language [1]. Using the visual world paradigm, previous work [2] has shown that verbs that imply fast or slow events can influence inspection of visual scenes, consistent with a simulation of the speed of the implied journey described in the language. For example, while concurrently viewing visual scenes featuring an agent on one side, a goal on the other side, and a path which extended between the two (see Figure 1), when hearing “The repairmen will race/hobble along the road to the car”, listeners looked at the path region longer after ‘hobble’ (slow) than ‘race’ (fast).

In the present work we describe two additional studies where we manipulated the speed of the event described in the language. In Experiment 1, speed of the event was varied depending on the verb, based on ratings collected in earlier work [2]. Participants listened to sentences, such as “The repairmen will race/hobble to the car along the road” whilst seeing accompanying visual scenes like Figure 1. In the new sentences the goal phrase was positioned before the path phrase in order to examine whether the constituent order (agent - > goal - > path) could affect mental ‘simulations’ manifested as eye movements on the visual scenes (agent - > path - > goal). As in previous work [2], results indicated evidence for increased dwelling on the path region of the scenes with slow verbs compared with the fast verbs, and crucially this was found before the path phrase (during ‘race/hobble to the car’).

In Experiment 2, we investigated two questions to extend the results of Experiment 1. We looked at (a) whether effects were specific to verbs, and (b) whether they were dependent on the presence of a prepositional path phrase in the sentence. Instead of verbs, we used adverbs to indicate differences in the implied speed of the sentences (e.g., fast – ‘quickly’, ‘hurriedly’, ‘energetically’; slow – ‘slowly’, ‘lazily’, ‘sluggishly’), which were paired with ‘speed-neutral’ motion verbs, e.g., “The repairmen will go hastily/gently to the car”. While we found a trend for increased dwelling on the path for slow event sentences (during ‘go hastily/gently to the car’), this was not statistically significant. However, in addition to eye tracking, we included a separate mouse tracking task as another way to index interpretation of the events described in the sentences. Participants heard the sentence while presented with a scene, and once the sentence ended, they had to use the mouse to ‘drag and drop’ the agent on the screen to the end location they thought was implied by the sentence. We examined the length of time it took for participants to move the agent, and here found clear evidence that it took longer to move the agent for the slow adverb sentences.

Our results overall are supportive of simulation accounts of meaning extraction during language comprehension [1], in which the length of time projected for an event to take place influences the amount of time looking along the path of the event, as part of a spatial updating process where language and scene information are dynamically integrated. Analogous results were found with motor movements. Our movement tracking results with adverbs indicate that evidence for simulations does not necessarily have to be tied to lexico-semantic information associated with verbs, or any other properties associated with those specific verbs, but instead can reflect comprehension of the event, given constraints integrated between language and the scene. However, the strength of these effects appeared to be influenced by lexically signalling the path of travel, which influences direction of attention to the scene.

ORTHOGRAPHY-SEMANTICS CONSISTENCY: A NEW MEASURE AND ITS IMPACT ON VISUAL WORD RECOGNITION
Marco Marelli (University of Trento), Simona Amenta (University of Milano-Bicocca) & Davide Crepaldi (University of Milano-Bicocca).
marco.marelli@unitn.it

A largely overlooked side result in most studies of morphological priming is a consistent main effect of semantic transparency across priming conditions. That is, participants are faster at recognizing stems from transparent sets (e.g., farm), in comparison to stems from opaque sets (e.g., fruit), regardless of the preceding prime types (e.g., farmer vs. fruitless vs. unrelated primes). This suggests that the semantic transparency of the derived form may be also consistently associated with some property of the stem word.

We propose that this property might be traced back to how well we can predict the meaning of a word on the basis of its orthography. WIDOW, to illustrate, has a very informative form, because it is related in meaning with any word in the English lexicon with the same onset (WIDOWER, WIDOWED, WIDOWHOOD). Conversely, WHISK is a less informative orthographic chunk because its original meaning is not retained in orthographically similar words (WHISKEY, WHISKY, WHISKERED, WHISKER). We name this property Orthography-Semantics Consistency (hence, OSC). In the present study, we exploit distributional semantic models to quantitatively characterize OSC, and test its effect on visual word identification relying on large-scale data taken from the British Lexicon Project (BLP, Keeulers et al., 2012).

We first confirm the reliability of the “stem transparency” effect by showing that it holds when testing 335 stems taken from previous masked priming studies on the BLP data (Experiment 1).

In Experiment 2, we proceed to compute OSC. For each item used in Experiment 1, we extract all the words with a similar onset, forming a family of “orthographic relatives”. Using distributional-semantics methods (Turney & Pantel, 2010), we automatically generate the semantic similarity between a stem and each of its orthographic relatives. We finally compute OSC as the frequency-weighted average semantic similarity. The resulting OSC measure has a significant effect in the expected direction: stems taken from transparent sets have significantly higher OSC than stems taken from opaque sets, and OSC has a facilitatory effect on BLP latencies in lexical decision.

Finally (Experiment 3), we show that OSC has a general effect in visual word recognition by testing it on 1821 words randomly selected from the stimuli included in the BLP database. Also when considering this large set of items, the effect of OSC on lexical decision latencies is significant. Moreover, the OSC measure is not significantly correlated with either frequency, family size or word length, suggesting that it captures a new and independent property of lexical representations.

In conclusion, results indicated that (a) the “stem transparency” effect is solid and reliable, insofar it holds in BLP lexical decision times (Experiment 1); (b) an unbalance in terms of OSC can account for it (Experiment 2); and (c) more generally, OSC explains variance in a large item sample from BLP, proving to be an effective predictor in visual word access (Experiment 3).

THE INFLUENCE OF LEXICAL ASSOCIATION ON SYNTACTIC ANALYSIS:
EYE MOVEMENT EVIDENCE

Anna Fiona Weiß (University of Marburg), Franziska Kretzschmar (University of Mainz, University of Marburg), Ina Bornkessel-Schlesewsky (University of Adelaide, University of Marburg) & Adrian Staub (University of Massachusetts, Amherst). fiona.weiss@staff.uni-marburg.de

A central question in research on sentence interpretation concerns when and how syntactic and semantic information interact. Recently, so-called semantic reversal anomalies have been used to address this question. These occur when potential arguments of a verb occur in reversed syntactic functions, as in “The hearty meal was devouring the kids” (Kim & Osterhout, 2005). In this example, the hearty meal is a likely theme of devouring, but a poor agent. There is ERP evidence that, at least in English, anomalies of this kind do not engender an N400 effect at the verb (a component which is sensitive to semantic processes), but instead elicit a P600, which is traditionally linked to syntactic processing. This finding has been interpreted by some as suggesting that the comprehender treats the hearty meal as the theme of devour, and that the P600 arises because the form of the verb is then treated as incorrect. This interpretation (for a different view, see e.g. Brouwer et al., 2012) suggests that a computation of the thematic relationship between subject and verb may precede a full syntactic analysis.

Here, we assess the effect of semantic reversal anomalies on eye movements in normal reading. There is previous eye movement evidence that thematic relationships may influence processing of ambiguous sentences only after the construction of an initial syntactic parse (e.g., Rayner et al., 1983). But to the best of our knowledge there has been no previous examination of reversal anomalies in natural reading.

The present experiment (N=56 native English speakers) manipulated, in a 2x2 design, whether the subject was a good agent for the verb (anomalous vs. non-anomalous) and the degree of association between the subject in the anomaly conditions and the verb:

a) Currently the book is writing the author after […]. (high association – anomaly)
b) Currently the book is fixing the author after […]. (low association – anomaly)
c) Currently the author is writing the book after […]. (high association – non anomaly)
d) Currently the author is fixing the book after […]. (low association – non anomaly)

In addition, we manipulated, between subjects, the difficulty of the comprehension questions that were asked, in order to assess possible influences of reading strategy.

On the verb (writing / fixing) and the object (the author / the book) we obtained the expected effect of anomaly, with longer reading times and more regressions in the anomalous sentences. However, the results also revealed a significant interaction between anomaly, association and region of the sentence. On the verb, the anomaly effect was reduced in high-association sentences in first pass and go-past time. By contrast, on the object region the anomaly effect was larger in high-association sentences. These results were not modulated by the task manipulation.

These findings show that a while a high degree of association between subject and verb does not reduce the effect of anomaly overall, it does change the temporal profile of this effect, with the anomaly effect shifted to a later point in time when subject and verb are strongly associated. Our findings thus confirm that the processing of syntactic information is not entirely independent of the computation of plausible thematic relationships. Crucially, this pattern distinguishes the reading profile for semantic reversal anomalies from other structures involving conflicts between syntax and semantics. The results are compatible with a model in which the outputs of the processing of different features (e.g. animacy, word order, lexical association) are integrated in a common analysis step where the different outputs are specifically weighted in order to find a final interpretation. This weighting may be language specific (cf. Bornkessel-Schlesewsky et al., 2011), but appears uniform across different task settings.
DOES A PREDICTION BENEFIT OUTLAST GRAMMATICAL VIOLATIONS?
A DOUBLE-VIOLATION ERP STUDY.

Dominik Freunberger (University of Salzburg)
dominik.freunberger@sbg.ac.at

Prediction in language comprehension has been shown to have beneficial effects on the processing of incoming information, be it in terms of facilitated lexical-semantic access/retrieval, easier integration, or – on a behavioural level – reduced reaction times (RTs) for predictable versus unpredictable words [1,2]. However, it is unclear how strongly predictable linguistic input is processed when grammatically erroneous, i.e. whether or not the prediction benefit outlasts or interacts with grammatical violations.

To test this, we used 60 German sentences with a verb that was highly predictable and morphosyntactically congruent (= infinite; 1A) or incongruent (= finite; 1B), and another 60 sentences with a verb that was unpredictable but still plausible and congruent (= infinite; 2A) or incongruent (= finite; 2B). The EEG of 23 German native speakers was recorded while they were reading the sentences silently word-by-word and asked to give an acceptability judgment after each sentence. ERPs were then calculated for the critical words (CW; underlined).

1 A/B Der Torhüter behauptet, dass der rutschige Ball einfach zu halten/ *hältst war. The goalkeeper claims that the slick ball easy to stop/ *stops was.
2 A/B Die Kinder prahlen, dass das junge Pferd einfach zu bürsten/ *bürstest war. The kids boast that the young horse easy to brush/ *brushes was.

Lower RTs for ungrammatical compared to grammatical sentences suggest that participants employed an error detection strategy to accomplish the task. However, for grammatical sentences, there seems to be a beneficial prediction effect reflected in the RTs, which is absent in the ungrammatical condition (RTs: 1B = 2B < 1A < 2A); yet, a ceiling effect as an alternative explanation cannot be excluded. Regarding the ERPs (figure) of the grammatical sentences (1A/2A), unpredictable verbs led to a typical biphasic N400-P600 effect relative to predictable verbs, indicating easier lexical-semantic access/retrieval (N400) as well as easier subsequent integration (P600) for the predictable verbs. Ungrammatical sentences (1B/2B) also led to an N400 effect for unpredictable versus predictable verbs, comparable to the N400 effect of the grammatical sentences. Additionally, both ungrammatical verbs led to a pronounced P600, but strikingly, without the positivity effect for unpredictable versus predictable verbs, like in the grammatical sentences, hence showing that these ERP effects are not additive. This might indicate that although lexical-semantic processing is facilitated for predictable vs. unpredictable words (N400), the subsequent integration, as reflected in the P600, benefits from greater predictability only, when the input is grammatical.

References
The Misinterpretation of Noncanonical Sentences Revisited
Markus Bader (Goethe University, Frankfurt) & Michael Meng (University of Applied Science, Merseburg)

The good-enough approach to interpretation (Ferreira, 2003) ascribes misinterpretations of noncanonical sentences to non-algorithmic, heuristic processes within the parser. However, whether interpretation errors occur on-line during parsing or off-line at a postinterpretive level is still an open question (Gibson, Bergen & Piantadosi, 2013). To address this question, we ran two comprehension experiments both combined with a reading span task. Experiment 1 adapted Ferreira (2003) to German. It tested the three syntactic structures in (1).

(1) a. SO: Der Koch hat den Braten ruiniert. ('The chef ruined the roast.')
   b. OS: Den Braten hat der Koch ruiniert. ('The roast, the chef ruined.')
   c. Passive: Der Braten wurde vom Koch ruiniert. ('The roast was ruined by the chef.')</n
72 triples consisting of two nouns and a verb were created. In condition Linking 1, noun 1 was the subject and noun 2 the object; in condition Linking 2, it was the reverse. Three types of semantic relations obtained between the nouns and the verb (see (2)): irreversible – linking 1 plausible and linking 2 impossible; biased – linking 1 more plausible than linking 2, but linking 2 still possible; symmetrical – linking 1 and linking 2 equally plausible. The results of a rating pretest are included in (2) (1=completely implausible; 7=completely plausible).

(2) Irreversible: linking 1: boy clean pan — 6.6; linking 2: pan clean boy — 1.2
    Biased: linking 1: chef ruin roast — 6.4; linking 2: roast ruin chef — 3.2
    Symmetrical: linking 1: father hug uncle — 6.1; linking 2: uncle hug father — 6.2

Experiment 1 used a procedure similar to Ferreira (2003), the major difference being that sentences in our experiment were presented visually, one word at a time in the center of the screen with a mean presentation time of 300ms per word. After sentence presentation, participants were cued to either name the actor or the undergoer.

The results for Experiment 1 are shown in Figure 1. The major findings of Ferreira (2003) are replicated. First, accuracy was highest for SO sentences, somewhat reduced for passive sentences, and even further reduced for OS sentences. Second, accuracy decreased with decreasing plausibility. Accuracy was highest for plausible irreversible sentences and lowest for implausible irreversible sentences, with the other sentence types in between.

In order to test whether the errors observed in Experiment 1 occurred on-line or off-line at a post-interpretative level, Experiment 2 presented the same sentences in the same speeded way as in Experiment 1, but required a simpler task. After reading a sentence, participants quickly had to decide whether the sentence was plausible or not. A correct plausibility judgment requires a faithful syntactic representation, but poses only few interpretative demands, especially for irreversible implausible sentences which can be rejected based on the selectional constraints imposed by the verb on its arguments. As shown in Figure 2, Experiment 2 yielded markedly different results than Experiment 1. In particular, irrespective of syntactic structure participants were highly accurate in rejecting implausible irreversible sentences—that is, the kind of sentences with the highest error rates in Experiment 1.

In sum, our data suggest that misinterpretations of noncanonical sentences arise at a post-interpretative level which follows the algorithmic processes of parsing and interpretation.
SEMANTIC INTERFERENCE EFFECTS IN PICTURE NAMING: THE EFFECT OF MANIPULATING TYPE OF SEMANTIC RELATION ON NAME RETRIEVAL TIMES

Melanie Vitkovitch, Anita Potton, Hisma Begum, Aimen Hasham and Mary-Maria Johnson (University of East London)

M.Vitkovitch@uel.ac.uk

There has been much work on exploring semantic interference effects in picture naming, with most work presenting pictures from well-established categories (e.g., animals, fruits, vehicles). Recent work by Abel Rahman and Melinger (e.g., 2011) has investigated different types of semantic relations between pictures, such as ad hoc categories. Pictures were presented in sequences related by theme (e.g., fishing trip, with pictures of stool, bucket, coffee), and repeated several times (blocking paradigm, with control condition). Semantic interference was only evident when an appropriate context title was given (e.g., “fishing trip”). They emphasized the dynamic nature of semantic access, with the context title promoting conceptual linkage between items.

In the first two experiments to be reported, a cumulative semantic interference (CSI) paradigm was used, and type of semantic relation between pictures was manipulated. Sets of 8 pictures were presented in sequence, drawn from either a) well-established semantic categories b) from different semantic categories but sharing a feature (e.g., made of wood) c) from different semantic categories, but related by goal/theme (e.g., fishing trip) and d) unrelated pictures. The sets of 8 pictures were randomized, as was the allocation of pictures to ordinal position within set (positions 1 to 8). There was no repetition of pictures. According to current theoretical accounts of the CSI effect, shared (semantic) activation is a key component underlying the effect. We expected a linear increase in naming times for well-established categories (CSI effect), where objects have a high degree of conceptual similarity, and reduced (if any) CSI effect for the semantic feature condition. The CSI effect would not be expected for objects related only by theme.

In Experiment 1, we replicated the CSI effect for well-established categories, but no linear increase in naming times was evident in the unrelated condition or semantic feature condition. In Experiment 2, which included also ad hoc categories, a contextual title was given for 2 seconds between each set of 8 pictures, with the aim of inducing interference effects between objects in the semantic conditions (cf Abdel Rahman & Melinger, 2011). Surprisingly, a quadratic effect was evident in the trend analysis, for all conditions. This suggested generic strategic effects resulting from the titles presented between the sets of 8 stimuli. There was some evidence in both experiments for overall slower response times in the semantic feature condition. In a final blocking experiment using contextual titles, we presented 6 pictures from each feature condition, for repeated naming (5 repetitions). Naming times were slower in the feature condition relative to the unrelated condition. From the results, there is no clear evidence that CSI effects can be found for semantic relations that are not “hard-wired”. However, given the results of the blocking experiment, and a growing interest in comparing these two paradigms, further work will be needed to establish whether CSI effects could emerge over repetitions.

LIMITS TO CROSS-MODAL SEMANTIC AND OBJECT SHAPE PRIMING IN SENTENCE CONTEXT

Joost Rommers (Max Planck Institute for Psycholinguistics) & Falk Huettig (Max Planck Institute for Psycholinguistics & Donders Institute for Brain, Cognition, and Behaviour)
joost.rommers@mpi.nl

Many studies have documented semantic priming effects from both words and pictures on word targets, but the literature on object shape priming in language processing is less well developed. Priming is typically observed with isolated words as targets. Some studies have shown that in sentence contexts, priming is not an automatic consequence of speech processing (Norris et al., Cognitive Psychology 2006). In addition, priming tasks tend to involve meta-linguistic judgments. In the present study we focused on cross-modal influences, which may occur when listening to spoken sentences while being situated in a visual environment. We tested effects of picture and written word primes on processing of target words embedded in sentences. The primes were related to the targets in meaning or object shape. We investigated whether these aspects automatically prime spoken-word processing even in sentence contexts and without a judgment task.

EEG was recorded from 23 adult native speakers who listened to spoken sentences and viewed written words and pictures. Each trial began with a short beep and a central fixation cross. A neutral sentence began playing (e.g., “She wrote a term paper about the eye”). 1000 ms before target word onset (“eye”), a prime appeared on the screen. The prime was either a written word or a picture. It was semantically related to the target (arm), had a similar visual shape to the target (ball), or was unrelated (javelin; n=32 items per condition, counterbalanced). After 500 ms, the prime disappeared and was replaced by the fixation cross. Participants were asked to carefully listen for comprehension while looking at the center of the screen. 64 word and picture filler primes which matched with the spoken words (e.g., a picture of a banana combined with a spoken sentence containing the word banana) were included to make the prime-target relationships less noticeable. To quantify the N400 component as an index of semantic processing, we averaged across a 300-500 ms time window and used a cluster-based permutation test to determine which electrodes showed differences. We also performed time-frequency analyses of oscillatory activity.

The ERPs to target words showed a clear N400 component in each condition. Within the written-word prime conditions, the N400 was attenuated in the shape condition relative to the semantic condition, whereas the other conditions did not differ from one another. Within the picture prime conditions, N400 amplitude did not differ reliably between the different prime-target relationships. Clear differences between picture and word primes during prime as well as spoken-target-word processing confirmed statistical power. The time-frequency analyses showed no effects of prime-target relationship during target word presentation, although during prime presentation alpha band power was lower for picture versus word primes.

The fact that prime-target relationship mattered for written word primes but not for picture primes suggest that there are limits to cross-modal priming. The present findings provide complementary evidence to previous studies on the automaticity of priming from speech, by showing that priming of speech processing is not an automatic consequence of seeing pictures.
BEFORE AND AFTER, AND PROCESSING PRESUPPOSITIONS IN DISCOURSE
Ming Xiang, Emily Hanink, Genna Vegh (University of Chicago)

mxiang@uchicago.edu

The connectives “before” and “after” order eventualities in the temporal domain. Münte, Schiltz and Kutas (1998, Nature) showed that, compared to after-initial clauses, before-initial clauses elicited a larger prolonged left frontal negativity, as early as 300ms after the onset of the connective (e.g. Before/After the scientist submitted the paper, the journal…). Together with results from other behavioral studies, comprehension cost on “before” has standardly been taken to show the influence from our conceptual knowledge about how events are temporally sequenced in the real world: more computations take place for before-initial clauses because they present events out of chronological order, whereas after-initial clauses do not. In the current study, we show that the original results arise not because of this clash in temporal sequence, instead, it is an example par excellence of our rapid sensitivity to presuppositions in discourse.

Temporal before and after-clauses generally presuppose the veridicality/factuality of the eventuality they denote. For instance, the default interpretation of (1) is that John indeed won an Oscar. However, although after-clauses are exclusively veridical, before-clauses allow non-veridical events, and receive an interpretation that the denoted events are likely to have had happened, as shown in (2), where changing “before” to “after” leads to contradiction (Heinämäki, 1974; Condoravdi, 2010). Comprehenders therefore couldn’t be certain whether the veridicality presupposition for before-clauses is satisfied until they read the main clause. The process of maintaining such uncertainty may be the source of the prolonged negativity observed for “before”, which is associated with extra working memory (WM) burden.

(1) Before/After John won the Oscar, he bought a big house.
(2) Before/*After the bomb exploded, it was defused by the police.

We conducted an ERP experiment to test this hypothesis, using two sets of stimuli. The first set contains “Ad-hoc events”, which are arbitrary events, as exemplified by (1) (and the original stimuli in Münte et al., 1998). In the “Real-World events” set, the temporal clauses always contain a true, widely-known, historical/cultural event, as in (3):

(3) Before/After Star Wars came out, George became interested in astronomy.

Since the “Real World” events eliminate any uncertainty associated with the veridicality of the before-clauses, we predict the absence of the negativity differences between before- and after-clauses, in contrast to the “Ad-hoc” events. A hypothesis based on conceptual temporal sequence predicts larger negativity on before-clauses under both types of events.

Experimental stimuli included 80 Ad-hoc and 80 Real-World before/after sentences, and an additional 80 filler sentences with when-clauses. ERPs were obtained over a prolonged 6200ms time window (with an additional 200ms pre-stimulus baseline) from the onset of the connective before/after, with the first 3150ms being the temporal clause, and the next 3050ms the main clause. Results from 30 native English speakers showed that over the 1000-6200ms time window there is a ROI (anterior to parietal distribution) x Event (Ad-hoc or Real-World) x Hemisphere interaction (F(2,58)=3.5, p<.05), and a marginal ROI x Event x Connective (Before or After) x Hemisphere interaction (p<.1). Further analyses revealed an Event x Connective x Clause x Hemisphere interaction (p<.05) in the anterior region (8 electrodes); crucially, although for Ad-hoc events, before-clauses elicited larger prolonged negativity than after-clauses (p<.01), no difference was found for before and after-clauses with Real-world events.
The aim of the study was to demonstrate that the parser can (i) use case marking to predict upcoming syntactic structure in non-verb-final languages (extending claims in Kamide et al. 2003), and (ii) project optional heads in the syntactic structure in order to maintain incremental processing (Sturt & Lombardo 2005).

We investigated sentences with ditransitive verbs in Russian. In such sentences when the verb is affirmative, the direct object is obligatorily Accusative and the indirect object is Dative. When the verb is negated, the direct object may either appear in the Accusative case (1a, a default scenario across languages), or in the Genitive case due to the existence of a Russian-specific ‘Genitive-of-negation’ dependency (1b). Whereas an Accusative case on the direct object is licensed by the verb (both in affirmative and negative sentences), a Genitive case on the direct object is licensed by Negation.

1a/1b. IvanNOM cvetyACC/cvetovGEN MašeDAT ne podaril. IvanNOM flowersACC / flowersGEN MashaDAT not gave ‘Ivan didn’t give flowers to Masha.’

Interestingly, the standard incremental parsing procedure for incorporating a pre-verbal direct object into the parsing tree, i.e., projecting a single verb head that licenses both the object’s thematic role and case, is inapplicable to Genitive-of-negation sentences (1b). Instead, the parser needs to predict multiple heads – the verb and (otherwise optional) Negation – on the basis of a preceding Genitive-marked object in order to incrementally incorporate that object. The study aimed to explore whether this is indeed the case.

Forty native speakers of Russian took part in a self-paced reading experiment in which they read sentences such as (2a-d). ([2a-d] are abridged versions of real sentences for space reasons.) A 2x2 design with factors Linear Order (Object-first vs Negation-first) and Case (Acc vs Gen direct object) was used. The critical conditions were Object-first sentences. If the parser is able to predict Negation (as well as the verb) on the basis of a linearly preceding Genitive-marked direct object, we expect faster reading times at the negated verb in (2b) vs (2a). Negation-first sentences were used to control for possible biases towards ACC vs GEN case on the direct object in negated sentences.

2a/2b: Direct Object-first
RoditelNOM reshili [gromozdkie podarkiACC/gromozdkix podarkovGEN buduščim molodoženamDAT ne darit’… ParentsNOM agreed [bulky giftsACC / bulky giftsGEN to newlyweds-to-beDAT not give

2c/2d Negation-first
RoditelNOM reshili [ne darit’ buduščim molodoženam gromozdkie podarkiACC/gromozdkix podarkovGEN … ParentsNOM agreed [not give to newlyweds-to-beDAT bulky giftsACC / bulky giftsGEN

‘The prudent parents have agreed to not give bulky items to the newlyweds-to-be ...’

There were no reading time (RT) differences at or following the direct object in (2c/d), suggesting no marked preference for Acc vs. Gen object in negative sentences. In contrast, there was a significant slowdown in RTs at the negated verb (not give) in the Accusative (2a) as compared to Genitive (2b) (RTs at the verb give in (2a): 390ms, (2b): 367ms., 95%CI [-44.6,-4.2]) Cumulatively, these results suggest that the parser uses case-marking on preverbal NPs to anticipate upcoming syntactic structure and can project multiple heads when a single head is insufficient to ensure incrementality of processing. Our findings extend previous claims on the use of case information for projecting the verb before it appears bottom-up in verb-final languages (Kamide et al., 2003; Miyamoto, 2002; Yamashita, 1997) to non-verb-final languages.
CASCADED LEXICAL PROCESSING TO SEMANTICS IN SPOKEN EMBEDDED WORDS

Jenna Todd Jones¹, Jeffrey Bowers¹, Nina Kazanina¹ (¹ University of Bristol)
j.toddjones@bristol.ac.uk

Previous research has revealed that visually presented subset words embedded within a carrier word are processed to the level of semantics and can interfere with performance in a semantic categorisation task (Bowers, Davis, & Hanley, 2005). For example, a participant is presented with a category (e.g. “Is this a vehicle?”) followed by words belonging (e.g. truck), or not belonging (e.g. tree), to that category. The critical words in the study did not belong to the stated category but contained embedded words that did belong to the category, e.g. the word scar (containing the embedded word car). The carrier word (scar) demands a ‘no’ response as to whether this word belonged to the vehicle category, while the subset word (car) would demand a ‘yes’ response. Bowers et al. found increased reaction times (RTs) and error rates to visually presented conflicting items (i.e. carrier words that contained embedded words belonging to the specified category) relative to non-conflicting items (when both the carrier word and the embedded words were unrelated to the category). This indicated some interference from the embedded word during semantic categorisation of the carrier, and was thought to indicate that lexical processing of the subset words had reached the level of semantics.

Method: In order to establish whether lexical activation of embedded words occurs in spoken words 30 participants were presented with 8 categories followed by 119 auditorily presented words requiring either a “yes” or “no” response as to whether they belonged to that category. As above, critical words were either conflicting (the carrier demanded a “no” response while the subset word demanded a “yes”) or non-conflicting (both the carrier and the subset word demanded a “no”). The embedded word occupied either an initial, middle, or final position within the carrier word, and crucially had either the same or different pronunciation as if it had been heard within the carrier word, e.g. ram within ramp (same) as opposed to ear within earn (different; see Table 1 for more stimuli examples).

Results: Conflicting stimuli yielded longer RTs (907ms versus 893ms, t(29)=3.95, p < .001), and higher error rates (1.6% versus .7%, t(29)=5.19, p < .001) than non-conflicting stimuli. RTs and error rates from conflicting trials (e.g. Is ‘ramp’ an animal?) were then subtracted from non-conflicting trials (e.g. Is ‘ram’ a vehicle?) to produce a conflict score that revealed the amount of semantic interference per stimulus type. These data were submitted to individual paired t-tests and revealed a significant difference in RTs and errors for only initially embedded subset words with the same pronunciation (see Table 1 for full statistics). Results from this paradigm provide evidence that when words are presented auditorily subset word processing cascades to the level of semantics to produce interference effects when attempting to semantically categorise the larger carrier word, but only in cases of onset-embedded words. We discuss our finding in the context of recent ERP research that has found effects for final-embedded spoken words, but only following a supporting sentential context (van Alphen & van Berkum, 2009).

Table 1. Stimuli examples and paired t-test results for conflict scores for auditorily presented embedded words (conflicting minus non-conflicting RTs and Errors).

<table>
<thead>
<tr>
<th>Embedding position</th>
<th>Pronunciation</th>
<th>Example item</th>
<th>RTs (ms)</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>Same</td>
<td>ramp</td>
<td>-35.1</td>
<td>&lt; .001*</td>
</tr>
<tr>
<td></td>
<td>Different</td>
<td>earn</td>
<td>-7.8</td>
<td>&gt; .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-12.2</td>
<td>&lt; .01*</td>
</tr>
<tr>
<td>Middle</td>
<td>Same</td>
<td>scowl</td>
<td>-15.4</td>
<td>&gt; .05</td>
</tr>
<tr>
<td></td>
<td>Different</td>
<td>crate</td>
<td>-5.9</td>
<td>&gt; .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Final</td>
<td>Same</td>
<td>howl</td>
<td>-8.9</td>
<td>&gt; .05</td>
</tr>
<tr>
<td></td>
<td>Different</td>
<td>warm</td>
<td>-10.7</td>
<td>&gt; .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.4</td>
<td>&gt; .05</td>
</tr>
</tbody>
</table>
FROM GUESSING TO GETTING IT RIGHT:
THE INITIAL DEVELOPMENT OF AN L2 LEXICON
Geertje van Bergen, Béryl Hilberink & Ton Dijkstra (CLS, Radboud University Nijmegen)
g.vanbergen@let.ru.nl

This study investigates L2 word recognition at the very first moment of contact with the L2, and during the 12 subsequent hours of L2 exposure. So far, only few studies based on natural language input have investigated this process at such an early stage of second language acquisition (e.g. Gullberg et al. 2010; Carroll 2012; Shoemaker & Rast 2013). Researchers involved in these studies disagree on the exact role of frequency in L2 word recognition: whereas some claim an important role of frequency in recognizing words in the early stages of L2 acquisition, others claim a limited frequency effect at this stage (cf. Carroll 2012). These diverging viewpoints seem to be partly due to the way in which frequency is measured, as well as its interaction with other factors, such as cognate status, which is known to facilitate L2 word recognition (e.g. Dijkstra et al. 1999).

The present study aims to get a better understanding of the role of frequency in the earliest stages of L2 word recognition. A 10-day Polish language course (approximately 1.5 hours of classroom instruction by a Polish native speaker daily) was designed especially for our research purposes. For each lexical item in the course, the intended frequency of use by the teacher was determined beforehand; also, we carefully controlled cognate status, i.e., phonological and semantic similarity between a lexical form in the L2 and the L1 (e.g. cognate - Polish professor vs. Dutch professor 'professor'; non-cognate - Polish tłumacz vs. Dutch vertaler 'translator'). All classroom interactions were recorded and transcribed afterwards, by which we were able to exactly determine the frequency of each lexical item at each point in time during the language course.

Dutch native speakers with no previous knowledge of Polish or any other Slavic language (N=40) took part in the language course. During the course, they were tested on their ability to distinguish Polish words from nonwords at five time intervals (after 0, 3, 6, 9, and 12 hours of classroom instruction) by means of an auditory lexical decision task. Stimuli were manipulated for Input frequency (number of occurrences in classroom) and Cognate Status (cognate vs. non-cognate); performance was assessed using a combination of behavioral measures (d', error rates and RTs).

Results indicate that initial learners very quickly move away from a guessing strategy and build up lexical representations in the L2: participants were significantly better and faster at distinguishing words from nonwords already after 3 hours of classroom instruction, and error rates and RTs diminished further across sessions. Within each session, items with a higher input frequency yielded more accurate and faster responses, although the frequency effect was restricted to non-cognates at later sessions. The effect size of input frequency on error rates and reaction times diminished as overall L2 exposure increased. We will discuss the implications of our findings for accounts of bilingual word recognition.
PUTTING THINGS IN NEW PLACES: VERB-BASED PREDICTION IN L1 AND L2 SENTENCE PROCESSING

Monique Flecken (Donders Institute, Radboud University) & Geertje van Bergen (CLS, Radboud University)
m.flecken@donders.ru.nl

Previous Visual World studies show that listeners anticipate sentence continuations on the basis of the available linguistic and visual context. In particular, verbs facilitate predictive processing (Altmann & Kamide 1999; Kamide & Altmann 2003). The present study investigates (1) whether language-specific verb semantics generates anticipatory eye movements in native listeners during incremental language understanding; (2) whether L2 users generate predictions on the basis of L2 semantic distinctions that are absent in their L1. We thereby focus on the semantic granularity of Dutch placement verbs: Whereas English uses the verb ‘to put’, Dutch distinguishes between zetten ‘to put-stand’ and leggen ‘to put-lay’, on the basis of the position of to-be-placed objects on a surface.

For 64 objects, prototypical positions were determined on the basis of a forced-choice sentence completion task, in which Dutch participants (N=30) filled in either staan (‘to stand’) or liggen (‘to lie’) to describe the position of objects (Experiment 1). In Experiment 2, participants (20 L1 Dutch, 20 L2 Dutch with L1 English) heard sentences describing placement events while viewing 4-object displays (192 trials) (‘look and listen’ task). We manipulated the factors Verb type (zetten, leggen or the neutral verb plaatsen ‘to place’) and Object position, both within objects (e.g., a standing vs. a lying bottle) and between objects (e.g., a standing frying pan vs. a lying ball) in each display. For each sentence (e.g., de jongen legde kort geleden een blikje op de tafel ‘the boy put-lay recently a can on the table’) the display contained (a) the target object in target position (a lying can), (b) a different object in target position (position-distracter, e.g. a lying fork), (c) the target object in a different position (object-distracter, i.e., a standing can), and (d) a different object in different position (distracter-control, e.g. a standing mirror), all placed on a table. We hypothesized that, upon hearing the placement verb, native speaker predictions should be biased towards objects whose position match the position implied by the verb, resulting in significantly fewer fixations on objects in non-target positions soon after verb onset. L2 users of Dutch should not generate predictions concerning sentence continuations on the basis of verb-specific semantics at all, or be slower in their predictions (cf. Martin et al. 2013).

Analyses show that native listeners are fast to narrow down possible sentence continuations (objects) to specific elements of the visual world. Factors that influence this process are (1) degree of match with the specific semantic granularity of information encoded in verbs, and (2) degree of conformity to prototypical object properties (prototypical positions, as established in Experiment 1. Furthermore, the L2 data show that predictive processing in this group is less automatized. We will discuss our findings in light of frameworks of prediction in incremental sentence processing. Overall, we support views of an incremental processor, using different sources of information (linguistic and ‘world-knowledge’ type of information) to reach an as rich as possible interpretation from early on (cf. Kamide et al. 2003). Furthermore, we will discuss how these findings are informative with respect to models of L2 processing.
Event-Related brain Potentials (ERPs) of semantic and grammatical processing have been investigated extensively, prototypically showing N400 modulations for semantic violations, and modulations of the P600 (or LPC) for certain grammar violations (e.g., agreement errors). Here, we focus on grammatical aspect, a core grammatical category with a strong meaning component, representing an interface between semantics and grammar. Progressive aspect in English marks an event as temporally ‘in progress’, and use depends on semantic constraints related to the temporal frame of reference introduced by discourse context (cf. Dahl 2000) (e.g., discourse contexts “What are you up to right now?” “I am working at home” vs. “What are you up to every Monday?” “I work at home”). Given its complexity, aspect is interesting for the investigation of second language (L2) processing. Here, we investigate whether L1 and highly proficient L2 readers of English are sensitive to aspect violations (a mismatch of discourse context and verbal aspect), and whether ERPs of aspect processing reflect semantic or grammar processing (Experiment 1). In Experiment 2, we obtained acceptability ratings of semantic, grammar and aspect violations from the same participants.

Participants (N=20 each group) read questions that set up a temporal-aspectual context (i.e., “What is Sophie doing in the pool right now?” progressive context vs. “What does Sophie do in the pool every Monday?” habitual context), followed by answer sentences. ERPs were time-locked to verb phrases in four conditions, e.g., progressive context followed by the answer “Right now, Sophie… (a) is swimming Control; (b) *is cooking Semantic violation; (c) *are swimming Grammar violation; (d) *swims Aspect violation; …in the pool” (same verb phrases rendered different conditions in habitual contexts).

Both groups showed typical N400 effects for semantic violations, and P600 modulations for grammatical violations. Aspect violations elicited a posterior N400 effect in the L2 group. In L1 readers we found an early, short Anterior Negativity (peak 250-300ms), and no N400 effect.

We will discuss individual differences in judgments of aspect violations and ERPs of aspect processing. Overall, findings are discussed in relation to studies identifying early Anterior Negativities as markers of predictive processing (Martin et al. 2013), reflecting a mismatch with expectations of temporal-aspectual reference frame generated by discourse context (see also Baggio 2008), potentially related to expectations at the level of orthography or form (Holcomb & Grainger 2006).
WHEN GROWING A FRUITLESS TREE IS A FRUITLESS EFFORT:
CONTEXT MODULATION OF THE PROCESSING OF MORPHEME MEANINGS IN COMPLEX WORDS
Simona Amenta (University of Milano-Bicocca), Marco Marelli (University of Trento) & Davide Crepaldi (University of Milano-Bicocca).
simona.amenta@unimib.it

How semantics inform morphological analysis at the early stages of visual word identification has been traditionally studied through masked priming, yielding results that are not completely uncontroversial. The majority of the studies suggest that semantics is not a factor during early morphological analysis, but there are also data showing that semantics do affect how we process morphemes early after a complex word has been seen (Davis & Rastle, 2010).

In a move towards more ecological testing conditions, in the present study we addressed this issue in natural sentence reading. In particular, we exploited a feature of several derived Italian words, that is, that they can be read in a “morphologically transparent” way (as genuine derivations) or in a “morphologically opaque” way (as pseudo-derivations) according to the sentence context where they belong. For example, the word copertina, which is made up of the stem copert-, blanket, and the productive diminutive suffix -ina, can either refer to a small blanket; thus being a fully transparent derived word (e.g., “Era una fredda sera d’inverno, così Lisa si mise sotto una copertina e accese il camino”, “It was a cold winter night, so Lisa covered herself with a small blanket and lighted up her fireplace”), or to a book/CD cover, which has clearly nothing to do with blankets (e.g., “Aprì il libro e, dopo aver letto la copertina, capì di avere trovato una prima edizione”, “She opened the book and, after reading the cover, she realized she had a first edition in her hands”).

31 of these words were embedded in sentence contexts that elicited either their transparent or opaque interpretation. By means of eye-tracking, we analyzed whether the effect of base word features – specifically, its context diversity (CD) – changes according to whether the (very same) word is read as a genuine derivation vs. as a pseudo-derived word, under the assumption that stem feature effects track morphological segmentation/analysis. Crucially, we focused on first fixations so as to make sure that we were looking at early stages of processing.

A mixed-model analysis revealed a base-word CD effect in both opaque and transparent contexts, thus showing that stem were accessed independently of whether they contributed to word meaning, that is, word decomposition is indeed blind to semantics. However, while the base-word CD effect was facilitatory in the transparent context, it was inhibitory in the opaque context, thus showing an early involvement of semantic representations. In other words, in the transparent context the word “copertina” is segmented into copert+ina and copert- act as a proper stem, thus facilitating the processing of the whole word. In contrast, in the opaque context, the word “copertina” is still segmented, but the pseudo-stem “copert-”, not contributing to the whole-word meaning, actually hinders the processing of the derived form. These data are in line with evidence on compound processing during sentence reading (e.g., Marelli & Luzzatti, 2012).

In conclusion, our data provide new and strong evidence that derived words are indeed segmented into their constituent morphemes when visually encountered, also in natural reading. However, this blind-to-semantics process activates morpheme representations that are semantically connoted, and their meanings can affect early stages of visual word recognition.

References
Structural Alignment Is Greater in Dialogue Than in Monologue

Iva Ivanova & Victor Ferreira (University of California, San Diego)
iva.m.ivanova@gmail.com

Interlocutors in a dialogue often mirror each other’s linguistic choices at different levels of linguistic representation, a process known as interactive alignment (Pickering & Garrod, 2004). According to these authors, alignment usually happens implicitly via an automatic priming mechanism hinging on a perception-action link, and can be affected by situational and partner-specific factors similarly to e.g. non-linguistic behavioural mimicry (see Chartrand & van Baaren, 2009). In a monologue setting, situational and partner-specific factors are either different or absent, hence alignment should function differently than in dialogue. Specifically, there is evidence that non-conscious behavioral mimicry such as foot shaking is enhanced when participants are provided with a goal to affiliate with each other (Lakin & Chartrand, 2003). If linguistic alignment hinges on a similar mechanism, it should occur more often in dialogue than in monologue (insofar as in the latter there is no reason to affiliate).

Here, we focus on structural alignment, and we report two experiments employing the structural priming paradigm, and comparing the magnitude of structural alignment in monologue and dialogue. Even though structural priming studies seem to report bigger effects in a dialogue (e.g., Branigan et al., 2000) than in a monologue setting (e.g., Ivanova et al., 2012), effects in the two types of setting have never been explicitly compared.

In both experiments, participants performed a picture-matching game. On each trial, they heard a dative prime sentence (e.g., The soldier gives the waitress the banana), then matched it to a subsequently presented picture, and then themselves described a dative-eliciting picture (e.g., a nun showing a boxer a ball). In Experiment 1, in one block (the solo block) the prime sentences were the pre-recorded picture descriptions of a research assistant, while in another block (the partner block) they were produced by another participant in the experiment, recruited in the same way as the “real participants” (those whose data was collected). Both the recorded and the other-participants’ descriptions were elicited by the same type of dative-eliciting pictures as used for the “real participants”, except they contained numbers indicating in what order the entities should be mentioned. The verbs were always the same between prime and target. Results revealed a significantly greater priming effect in the partner block (55%) than in the solo block (42%) (a difference of 13%).

Experiment 2 had a similar set-up, but the prime sentences in both blocks were produced by one of two experimenters, who read scripted sentences instead of describing pictures. Results revealed a non-significant difference in the priming effect between the partner (38%) and solo blocks (34%) (a difference of 4%). These results might differ from those of Experiment 1 because the interlocutors in the partner block (who were also the experimenters) might not have appeared as engaged in the task, due to performing it multiple times. Also, the prime sentences they read might not have had the same prosodic contours as naturalistic picture descriptions.

These results generally support Pickering and Garrod’s (2004) proposal that the alignment mechanism is sensitive to situational and partner-specific cues. They, however, suggest that it is not the sheer presence or absence of an interlocutor which influences alignment. Instead, the magnitude of alignment may be sensitive to subtler cues such as the interlocutor’s engagement in the task and/ or the prosodic contours of their utterances (see Kullen & Brennan, 2012), both of which might influence processing depth (Branigan et al., 2006). However, the possibility remains (and will be addressed in future work) that not one but two distinct mechanisms are at play during structural alignment – one that comes via priming of linguistic (here, structural) representations, and another hinging on imitation (akin to the non-linguistic behavioural mimicry effects) that takes into account situational and partner-specific cues to enhance or reduce alignment.
Natural languages share certain properties that presumably arise as a result of language learning and use (Christiansen & Chater, 2008). One such universal feature is the lack of unpredictable variation, when an object or grammatical marker is labeled by more than one morpheme or word that are used completely interchangeably. Moreover, learners receiving variable linguistic input tend to eliminate it, making the language more regular (Hudson Kam & Newport 2005). This makes unpredictable variation a good test case to study learners’ biases and how these biases give rise to regular languages through learning, transmission (Reali & Griffiths, 2009, Smith & Wonnacott, 2010) and use (Pickering & Garrod 2004). We explore how regularisation is influenced by social cues in the learners’ input language, in particular when variability is distributed within and across speakers. We presented participants with objects on a screen while playing them descriptions of the objects in an artificial language, and subsequently tested their ability to label the same objects verbally using the language they learned. The target language contained lexical variability: each object was associated with two labels. We manipulated how the variability was distributed across speakers of the language: learners either received input from one or three variable speakers, or from three speakers who were individually consistent but exhibited variability collectively (i.e. they used different labels). The input proportion between the two alternative labels was kept constant in all conditions in a 2:1 ratio.

For learners who learned from variable speakers, we found no significant difference between learning from a single speaker or from multiple speakers. However, there was a significant effect of variability in the language of multiple speakers. Participants who learned from variable speakers (Within Speaker variation) produced variable languages themselves: they used both labels in their output (Figure 1, red bars). On the other hand, participants learning from speakers who were individually consistent (Between Speaker variation) regularized more, and in their output were disproportionately likely to use the majority label - that used by two of the three speakers in their input (Figure 1, green bars).

This indicates that the propensity of language learners to eliminate linguistic variation is modulated by social cues, either by the process of conformist copying of the majority behavior or by copying the linguistic preferences (specifically, the degree of within-individual variation) of the speakers they learn from. These findings, particularly the potential conformity effects we observe, link the literature on the learning of unpredictable linguistic variation (and linguistic variation more generally) to the broader literature on social learning in other domains.

Figure 1. Proportion of objects (y axis) labeled with a given proportion of the Majority Label (x axis, bins) in the Within Speakers (red bars) and Between Speakers (green bars) conditions. Intervals include lower boundary but not upper boundary, except for rightmost bin.
Semantic properties of words, such as imageability or concreteness, affect their acquisition (Morrison, Chappell, & Ellis, 1997). Semantic properties of content open-class elements may also affect the acquisition of grammatical morphemes that combine with these elements (e.g. Bloom, Merkin, & Wootten, 1982). Recent research showed that semantic properties of words, such as imageability, can facilitate the retrieval and recognition of inflected forms in adults (Prado, Ullman, 2009). Such semantic effects are surprising because they are indirect: properties of a word determine how easily this word combines with inflectional morphemes. At the same time, the existing findings and suggestions provide theoretical and empirical motivation for examining the effects of semantic factors on the acquisition of grammatical morphemes. Such effects have indeed been documented in one previous study (Smolik, 2014). The semantic factor under investigation was imageability, which is known to affect children’s acquisition of vocabulary (McDonough et al., 2011).

The present study examined the effects of imageability on the acquisition of inflected forms while controlling for the age of acquisition of the uninflected forms. The study used a parent report questionnaire and examined Czech children aged 1.5 to 3.5. The forms examined were nominative plurals in nouns, and second-person singular present forms and past participles in verbs. The use of unmarked forms - nominative singular nouns and third person singular present verbs - was assessed. Total of 317 parents were asked to mark whether their child used the forms of interest in a total of 105 words (62 nouns and 43 verbs). The analysis used mixed-effects logistic regression with crossed random effects for words and persons, and estimated the likelihood of reporting that the child used the inflected form. The predictors included in the model were the use of the uninflected form, the child's age, the frequency of the inflected form, and the imageability rating of the word. The analysis found significant facilitative effects of imageability for all three inflected forms.

Together with previous research, the study provides converging evidence that semantic factors are involved in the acquisition of inflected forms. Two mechanisms may be responsible for this effect. Highly imageable words might be easier for lexical processing, thus leaving more resources for processing and acquiring the inflected forms. Or, highly imageable words may be easier to retrieve and that this effect extends to inflected forms. Possible methods for distinguishing these two mechanisms will be discussed.

References
Alphabetic languages vary considerably in their orthographic transparency, i.e. how easy it is to derive the pronunciation of a word from its spelling. English is opaque (e.g., in “ball”, the letter “a” is actually pronounced as /o/) while Finnish has one of the most transparent orthographies in the world. Like Latin, virtually all words are pronounced as they are written. The German orthography is generally considered to have an intermediate degree of transparency (e.g., the letter “a” is always pronounced as /a/ but its length is underspecified). Orthographic transparency has been shown to be an important factor in adults’ lexical processing and is argued to be the main determinant for cross-linguistic differences in reading development (e.g., Ziegler et al., 2010). Children for whom it is easier to “crack” the orthographic code are expected to learn reading earlier and to rely more on fine-grained lexical decoding strategies. However, this has never been investigated in silent reading using eye-tracking methods.

Here, we present first results from a study comparing eye-movements between children and adults in Finnish, German, and English (n = 30 and n = 25 in each language, respectively). Identical stimulus materials were used that comprised 96 sentences and over 800 words. In addition, word length and frequency were manipulated orthogonally for specific target words. Sentences were matched by structural and lexical aspects as far as possible and carefully translated and back-translated into all three languages. We are thus able to investigate the impact of orthographic transparency while holding semantic content constant.

In the present talk, we concentrate on length effects (Joseph et al., 2009), i.e., the impact of the number of letters on different eye-movement measures, which is a marker effect for sublexical decoding. Using linear (additive) mixed-effects models we analyze the form and strength of these effects in children and adults and compare them between languages. For adults, results showed that length effects varied strongly with orthographic transparency: They were rather strong in Finnish, intermediate in German, and weak in English. For children, by contrast, length effects were similar and equally strong in all three languages. Thus, developmental changes were particularly large in English, indicating a qualitative change from sublexical to lexical processing. In German and Finnish, developmental differences were merely quantitative and indicate that sublexical decoding remains a dominant processing strategy even in adults. Overall, our results show that children’s eye-movements continuously adapt to the constraints of their linguistic environment. Implications for models of eye-movement control are discussed.

LIMITS AND VARIATIONS OF LINGUISTIC GENERALIZATIONS

Silvia Radulescu (Utrecht University), Frank Wijnen (Utrecht University) & Sergey Avrutin (Utrecht University)
sil.radulescu@gmail.com

When confronted with the challenge of learning their native language, children manage impressively fast to infer generalized rules from a limited set of linguistic items, and apply those rules to strings of words never heard before. This study investigates what triggers and what limits the inductive leap from memorizing specific items to extracting general rules. Our new entropy-based approach allows the prediction that generalization is a cognitive mechanism that results from the interaction of linguistic input complexity (entropy) and the limited processing and memory capacity of the human brain (i.e. a limited channel capacity).

It has been argued that children detect patterns in auditory input, like phonotactic information (Chambers, Onishi & Fisher, 2003), and word boundaries (Saffran, Aslin & Newport, 1996), by statistical learning. Given that statistical learning deals with computing the probability that a sound in the input occurs after another sound, it cannot account for abstraction of rules beyond the input. A conceptual distinction was drawn between the types of abstractions made by learners (Gómez & Gerken, 2000): abstractions based on perceptual characteristics of specific elements (ba follows ba) and category-based abstractions (relations between abstract categories, e.g. Noun-Verb-Adverb). It was also proposed that humans have an algebra-like system (Marcus, Vijayan, Rao & Vishton, 1999), for extraction of general rules that apply to categories of items, such as “the first item is the same as the third item” (li na li). An algebra-like system addresses the case of generalizing to novel input, but it does not explain how humans tune into such algebraic rules, and what the factors (if any) in the linguistic input are that facilitate or impede this process. Our entropy model addresses these questions and bridges the gap between previous findings unifying them under one consistent account.

According to this model, less complexity in the linguistic input allows memorization of specific items, while a higher input complexity that overloads the channel capacity drives the tendency to make generalizations (i.e. reduce the number of features that individual items can be coded for and group them in abstract categories, and acquire relations between these categories).

We exposed adult participants to 3-syllable AAB strings that implemented a miniature artificial grammar to probe the effect of input complexity on the process of generalization. To obtain three degrees of input complexity, two factors were manipulated: the number of (different) syllables and the number of repetitions of each syllable. Entropy (a function of the number of items and their probability of occurrence) was used as a measure of input complexity, to design three experimental conditions: low entropy (4x6 A-syllables/4x6 B-syllables), medium entropy (2x12 As/2x12 Bs) and high entropy (1x24 As/1x24 Bs). Participants were asked to give grammaticality judgments on four types of test strings: grammatical AAB strings with trained syllables, grammatical AAB strings with novel syllables, ungrammatical novel ABC strings (three different new syllables), and ungrammatical ABC strings with trained syllables. The results showed that the higher the input complexity, the higher the tendency to abstract away from specific items and make a category-based generalization (i.e. consider novel strings with AAB structure to be grammatical). There was a roughly U-shaped performance for ungrammatical ABC strings with trained syllables: participants in the low entropy condition correctly rejected these test strings significantly more often than participants in the medium entropy condition, but only slightly better than participants in the high entropy condition. These results point to the interplay between the two tendencies - perceptually-bound learning and category-based abstraction – which were working against each other, pushing in different directions depending on their gained strength.

This entropy model can account for the inductive leap from perceptually-bound learning that memorizes and produces constructions with items encountered in the input, towards a category-based mechanism that applies abstract rules productively, in response to the degree of complexity (entropy) in the environment. Unlike previous findings this model also gives a quantitative measure for the likelihood of making generalizations in different ranges of input complexity.
THE BOUBA EFFECT: SOUND-SHAPE ICONICITY IS BASED ON ROUNDNESS
Matthew Jones (University College London), David Vinson (University College London), Nourane Clostre (University College London), Alex Lau Zhu (University College London), Julio Santiago (Universidad de Granada), and Gabriella Vigliocco (University College London). john.jones.09@ucl.ac.uk

Although wordforms are often arbitrarily linked to their meaning, many exhibit iconicity (resemblance between form and meaning). This is especially visible in the lexica of non-Indo-European languages and signed languages (Perniss, Thompson, & Vigliocco, 2010). Iconicity has been argued to play a role in grounding linguistic form to real-world experience, rendering language more learnable. Here we examine sound-shape iconicity, the 'kiki-bouba' effect, i.e. the tendency to associate bouba-like labels with round shapes, and kiki-like labels with spiky shapes (Ramachandran & Hubbard, 2001).

The study establishes the robustness of sound-shape iconicity by eliciting it in several new paradigms. It also disentangles the separate contributions of round and spiky sound-shape associations. These are impossible to tease apart in the classic kiki-bouba experiment, but if one were to have primacy this fact could have important implications for the origins of the effect.

In Experiment 1 we show that sound-shape iconicity appears in novel names people generate for round and spiky shapes.

Experiment 2 uses iterated learning (Kirby, Cornish, & Smith, 2008) – a model of cultural evolution. A first participant is taught an arbitrary 'language' of names for videos of moving shapes, and then tested on what she’s learned. The output from the testing goes to a second participant (along with any mistakes made), whose output goes to a third participant, and so on for ten participants (or generations). The 'language' changes over the generations, tending to retain any chance modifications that make it more learnable. We show that sound-shape iconicity emerges in the course of iterated learning chains (presumably because it renders labels more learnable). However, it only emerges for the mapping between round shapes and bouba-type labels.

Experiment 3 uses cross-situational learning (an implicit learning paradigm argued to mirror children’s exposure to new vocabulary - see Monaghan, Mattock, & Walker, 2012) to examine mappings of the bouba-to-round type and the kiki-to-spiky type separately. Bouba-to-round mappings aid learnability, but kiki-to-spiky mappings do not.

We will discuss possible mechanisms underlying this asymmetry. Ramachandran & Hubbard (2001) suggest that the kiki-bouba effect may be the consequence of metaphorical similarities between e.g. spiky shapes and articulatory gestures responsible for 'spiky' sounds; or alternatively the product of direct cross-modal mappings in brain architecture. However, these accounts at best accommodate an asymmetry, they do not predict it. Instead we conclude that the asymmetry could imply that the effect is based on visual or motoric representations of the round mouth shape required to articulate certain 'round'-sounding phonemes (e.g. rounded vowels).

References


SENSITIVITY TO PROSODIC CUES AND THEMATIC EMPHASIS IN CHILDREN’S SENTENCE CONTINUATION

Mirta Vernice (University of Milano-Bicocca) & Maria Teresa Guasti (University of Milano-Bicocca).

mirta.vernice@unimib.it

In connected utterances, the choice of which referent to mention as the subject of a subsequent sentence depends on which element appears as most salient and thus most accessible to the speaker (Arnold, 2008). Several linguistic cues might make an element salient, such as animacy, syntactic and linear position, prosodic stress or thematic emphasis (Bernolet et al., 2009). The literature on language acquisition claims that children process these different sources of information in the elaboration of the sentence, but they lack the ability to integrate them to assign an interpretation to the sentence (Trueswell et al., 1999). To address this issue, in the present study, we investigated whether prosodic and thematic emphasis cues are exploited by children to decide which referent is more salient in a utterance and thus more likely to be assigned to subject position in a subsequent sentence.

To determine to what extent Italian speakers make use of prosodic emphasis in order to decide which referent is more salient, we asked 4-, 5-, 6-year-olds (12 participants for each age-group), and adult speakers of Italian to generate a continuation of three types of sentences with the same surface order: (1) with default stress on the last NP (la zebra bagna l’ippopotamo; ‘the zebra is soaking the hippo’), as assigned by the nuclear stress rule (cf. Chomsky & Halle, 1968); (2) prosodic emphasis on the subject (la ZEBRA bagna l’ippopotamo; ‘the ZEBRA is soaking the hippo’); (3) prosodic emphasis on the direct object (la zebra bagna l’IPPOPOTAMO; ‘the zebra is soaking the HIPPO’). The dependent variable was the proportions of continuations where the object (i.e., hippo) appeared in subject position, i.e., object-next-mentions. 5-year-olds produced 12% more object-next-mentions after (3) as compared to (1). This tendency was consistent in 6-year-olds and adults too (respectively 25% and 13% more). Conversely, 4-year-olds did not appear to rely on prosodic emphasis cue in order to decide which referent to mention next (only 3% more after (3) as compared to (1)). That is, children up to four appeared to rely mostly on structural cues (syntactic and/or linear position) to determine which referent was most salient.

We ran a second experiment in which we investigated at what age children were sensitive to another source of emphasis, namely thematic emphasis, in order to determine the most salient element. To do so, we tested 3-, 4- and 5-year-olds Italian participants. The experimental set of sentences involved: (1) a matrix SVO (la zebra bagna l’ippopotamo; ‘the zebra is soaking the hippo’); (2) an It-cleft (E’ l’ippopotamo che la zebra sta bagnando; ‘it is the hippo that the zebra is soaking’) and (3) a pseudo-cleft sentence (Quello che la zebra sta bagnando è l’ippopotamo; ‘the one that the zebra is soaking is the hippo’).

3-year-olds tended to mention next, as the subject of their continuations the element that appeared in sentence initial position, i.e., after (2) there were 15% more object-next-mentions as compared to (3). Similarly, 4-year-olds produced 12% more object-next-mentions after (2) as compared to (3). Only 5-year-olds were sensitive to thematic emphasis, with only 4% more object-next-mentions after (2) as compared to (3). Importantly, after both (2) and (3), there were respectively 13% and 9% more object-next-mentions in comparison with (1).

The present results show that sensitivity to prosodic and thematic emphasis is not exploited verbally before five years of age. These findings suggest that a developmental change in the integration of information from different linguistic sources occurs between 4 and 5 years of age.
Subject-verb agreement attraction in production by Japanese learners of English
Mayuko Yusa (Tohoku University), Jungho Kim (Tohoku University), Noriaki Yusa (Miyagi Gakuin Women’s University) & Masatoshi Koizumi (Tohoku University)
mayuko.yusa@gmail.com

**Background:** It is well-attested that English native speakers sometimes make subject-verb agreement errors when the subject consists of a complex NP with the head noun followed by a modifier containing another noun, producing an ungrammatical sentence such as “The key to the doors are...” (Bock & Miller, 1991). In this case subject-verb agreement is interfered by the local or intervening noun doors. The phenomenon traditionally referred to as “agreement attraction” has attracted considerable attention in first language (L1) studies, which have thus far identified two cases of agreement attraction: (i) agreement attraction occurs when the head noun and the local noun disagree in number as in “The key to the doors” and “The keys to the door”; (ii) plural local nouns (e.g., “The key to the doors are...”) induce more agreement attraction errors than singular local nouns (e.g., “The keys to the door is...”). To date, almost all studies on subject-verb agreement attraction in second language (L2) have focused on comprehension experiments with participants whose L1s have overt subject-verb agreement. Therefore, the nature of agreement attraction in production by participants like Japanese learners of English whose L1 lacks overt subject-verb agreement still remains underspecified.

**The experiment:** In order to identify the nature of agreement attraction errors in L2 production we carried out sentence completion tasks in two experiments with 116 adult Japanese learners of English at an intermediate level (Oxford Quick Placement Test, 2001). The participants were visually presented with a sentence preamble and then were asked to complete the sentence preamble in order to produce a full sentence involving a copular verb BE. Experiment 1 tested whether Japanese L2 learners would be sensitive to hierarchical structure: the head noun followed by a complement PP as in “The drawing(s) of the flower(s)” vs. the head noun followed by an adjunct PP as in “The drawing(s) with the flower(s).” Experiment 2 investigated whether types of prepositions in an adjunct PP would affect agreement attraction errors: for as in “The assistant(s) for the inspector(s)” vs. with as in “The assistant(s) with the inspector(s).” In both experiments the number-combination of the head noun and the local noun was counterbalanced: the singular head noun followed by the singular local noun (SgSg), the singular head noun by the plural local noun (SgPl), the plural head noun by the singular local noun (PlSg), and the plural head noun by the plural local noun (PlPl).

**Results and discussion:** In Experiment 1, the participants produced significantly more agreement errors with the subject containing the complement PP than the adjunct PP (p<.05). In Experiment 2, no significant difference was found between for and with. These results suggest that participants are sensitive to syntactic structures and that agreement is produced on a hierarchical rather than linear structure. This argues against the Shallow Structure Hypothesis (Clahsen & Felser, 2006) that L2 learners are blind to detailed syntactic information. Interestingly, main effects for the number-combination (SgSg<PlPl=SgPl<PlSg) were found in both Experiment 1 and Experiment 2, where “X < Y” indicates that X induced significantly fewer agreement errors than Y (p<.001). The results (SgPl<PlSg) are the reverse of the L1 studies showing that native speakers make more agreement errors in the SgPl condition than in the PlSg condition (PlSg<SgPl). The number-combination asymmetry with native speakers has been taken as providing support for the feature-markedness theory that a number feature is marked on plural nouns but not on singular nouns: in the SgPl condition the plural feature on the plural local noun erroneously percolates to the whole subject, inducing agreement attraction errors; in the PlSg condition the singular local noun has no number feature to percolate, resulting in the correct agreement between the plural head noun and the verb (Eberhard, 1997). However, this theory cannot explain that Japanese L2 learners make more agreement errors in the PlSg condition than in the SgPl condition. We argue that L1 transfer will answer this problem, by referring to Chierchia’s (1998) Nominal Mapping Parameter, according to which Japanese nouns are mass nouns.
The effect of syntactic persistence, the unconscious repetition of sentence structures between utterances and speakers, provides evidence for the psychological reality of abstract syntactic representations in adults and young children (Bock, 1986; Huttenlocher et al., 2004). It has also been suggested that such representations are dynamic and susceptible to change caused by experience. As such, priming effects reflect implicit learning: Processing the syntax of a prime leads to longer-lasting changes in sentence processing than just immediate syntactic persistence (Chang et al., 2012). Support for this claim comes from studies showing long-term priming effects in adults (Bock & Griffin, 2000). The mechanisms of such learning effects in adults have been linked to the mechanisms of children's language learning. This suggests that syntactic priming effects should also persist in children but little attention has been paid to the longevity of syntactic priming in children versus adults (cf. Savage et al., 2006; Kidd, 2012). We examined whether children indeed show long-term priming effects and we compared this effect in children and adults.

The structure of interest was the passive: Whilst children have acquired some knowledge of the passive by three years and are susceptible to priming of passives, this structure is not frequently used in spoken English and children display difficulty comprehending and producing it until relatively late on (Messenger et al., 2011). Since actives are strongly preferred over passives, we would not expect children to use passives if not primed. Furthermore, less-preferred structures such as passives are more able to be strengthened by experience than frequently processed structures and so are more susceptible to learning effects. Passives thus provide an ideal test case for examining learning effects in children.

Method: We tested 57 pre-school children (M= 4.3yrs) and 64 adults. Participants described pictures of transitive events in a two-part turn-taking task. In part 1, Prime participants described 12 targets after hearing prime descriptions; participants heard both active (6) and passive (6) primes to avoid reinforcing the use of one structure only. In part 2, they described 12 targets after hearing syntactically-unrelated (adjectival) descriptions. Control participants described the same targets after hearing syntactically-unrelated, adjectival descriptions in both parts 1 and 2.

Predictions: Following exposure via passive primes, Prime participants should produce more passives in part 1 than Control participants (as well as more passives following passive primes than following active primes). We further predicted that if syntactic priming effects are long-lasting, Prime participants should also produce more passives than Controls in part 2.

Part One Results: Prime participants exhibited a robust priming effect (37%), producing more passives following passive primes than active primes, which did not differ significantly between children (42%) and adults (33%). Critically, Prime participants produced more passives than Controls; this difference occurred within both age groups but was greater in children than adults. Part Two Results: Overall, Prime participants produced more passives than Controls but this difference was only significant in children.

<table>
<thead>
<tr>
<th>Group</th>
<th>Part 1 : Mean number of passives</th>
<th>Part 2 : Mean number of passives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prime</td>
<td>2.65 (Children: 2.77; Adults: 2.53)</td>
<td>1.95 (Children: 1.80; Adults: 2.06)</td>
</tr>
<tr>
<td>Control</td>
<td>0.24 (Children: 0.03; Adults: 0.44)</td>
<td>0.60 (Children: 0.10; Adults: 1.09)</td>
</tr>
</tbody>
</table>

Conclusions: Participants who heard passive descriptions produced more passives than those who did not, even once they ceased to hear them. Children were more susceptible to this effect than adults. We conclude that priming effects persist beyond the immediate exposure and may be consistent with implicit learning mechanisms.
I present a model that investigates in a bottom-up way of how infants learn to imitate the speech sounds of a language. Infants’ aptitude for imitation is not self-evident: it has been shown that in infant-parent interactions, at least 2 to 6 month old infants imitated their parents’ vocalizations rarely, whereas the parents imitated the children’s vocalizations significantly more [1]. One problem is that the differences in vocal tract sizes and shapes between infants and adults make it impossible for an infant to produce the acoustic characteristics defining (for example) a native vowel to an adult. In order to succeed in imitation, infants have to learn to map acoustic features of their caregivers’ speech onto their own articulatory and acoustic domains. Infants also have to somehow explore their articulatory domain and converge onto speech sounds characteristic to their native language. Parents’ feedback on infants’ babbling has been shown to have an effect on the quality of babbling [2] but details of the learning process are still largely unknown.

Computational speech acquisition simulations offer a way to study potential mechanisms for learning speech; different models can easily be tried out, since virtual parent-child interactions can be run extremely quickly compared to real-life situations. Our aim is to find answers to what kind of rules or cognitive abilities are needed to make infants’ speech learning possible. We demonstrate the feasibility of the found mechanisms by creating a virtual infant - equipped with an articulatory synthesizer - that is able to imitate human vowels after a number of interactions with a human speaker. In contrast to using a virtual caregiver (see [3]), the current work uses a real human participant in teaching the virtual infant.

In the experiments, the infant babbles open vowel sounds pseudo-randomly. Based on observed behavior of caregivers and real infants, the participant is set to answer the babbles with real Finnish words with partially matching phonemic content. The infant creates auditory perceptual categories based on its own babbles and associates acoustic features of the heard answers to these categories. The infant’s imitation ability is tested using a new set of test words spoken by the participant, whose vowels the infant is set to imitate based on the learned associations and stored articulations related to the perceptual categories. Important factors to enable efficient learning include: 1) increasing articulatory accuracy due to babbling experience in certain articulatory regions, 2) a bias to try to either explore novel articulatory regions or repeat already learned perceptual categories, 3) active listening of caregiver’s speech and weighting babbling to those perceptual categories that are activated based on learned associations thus far, 4) clustering the infant’s learned phonemic categories based on their similarity to associations with the caregiver’s speech (which frees resources to explore more interesting categories when it seems likely that more than one of the infant’s categories are associated with one caregiver’s category). The pilot experiment shows that the current methods lead to over 90 % imitation accuracy for Finnish vowels in around 1000 babble-imitation pairs.

Tracking Emerging Sensitivity To L2 Morphosyntax: Evidence For The Role Of The L1 From Two Longitudinal Studies Of Early Learners

Alison Gabriele (University of Kansas), Robert Fiorentino (University of Kansas), José Alemán Bañón (University of Reading), Lauren Covey (University of Kansas) & Amy Rossomondo (University of Kansas)
gabriele@ku.edu

Several proposals on second language acquisition in the ERP literature argue for a similar starting point for the processing of all morphosyntactic structures: at the initial stages, all morphosyntactic violations should elicit either no effects (Steinhauer et al., 2009; McLaughlin et al., 2010) or an N400 (McLaughlin et al., 2010; Ullman, 2005), a component argued to index the strength of lexical associations (e.g. Kutas and Hillyard, 1980). At later stages, different structures may transition to the P600, a response argued to index syntactic analysis and repair (Osterhout and Holcomb, 1992), at different rates. In contrast, theories of transfer in L2 acquisition predict different starting points for different linguistic properties depending on the inventory of features in the L1 and L2 (e.g. Schwartz and Sprouse, 1996). Under the transfer proposals, features that are shared by the L1 and L2 may show evidence of a P600, a native-like response, from the earliest stages of acquisition.

The present study examines the role of the L1 in the processing of gender and number agreement in English-speaking learners of Spanish, tracking development in two longitudinal studies. In Study 1, first-year learners of Spanish were tested in three sessions after two months, six months, and eight months of exposure/instruction (n=23). In Study 2, a group of second-year learners of Spanish were tested in two sessions, after two months and six months of instruction in second-year Spanish (n=10, data collection ongoing). The experiment targeted three types of agreement: number agreement on verbs, which is similar in Spanish and English (1); number agreement on adjectives, which is a shared feature of the two languages, although instantiated differently in Spanish (2a,b), and gender agreement on adjectives, which is unique to Spanish (2a,c).

(1) Subject-Verb Agreement
La pasajera \textit{SG} brasileña (a) desembarca \textit{SG} / (b) *desembarcan \textit{PL} en San Diego.
‘The Brazilian passenger disembarks at San Diego.’

(2) Noun-Adjective Number/Gender Agreement
La biblioteca \textit{FEM.SG} es (a) moderna \textit{FEM.SG} / (b) *modernas \textit{FEM.PL} / (c) *moderno \textit{MASC.SG} y …..
‘The library is modern and …..(the school too).’

Results for the Spanish native controls (n=12) revealed reliable P600s for all violation types. For the first-year learners (Study 1), a small positivity emerged in midline electrodes for number violations on both verbs (1b) and adjectives (2b) across sessions. However, after sessions two (six months) and three (eight months), this positivity evolved into a more broadly distributed, canonical P600, which was significant at both the midline and lateral regions, although the effect was only numerical for number violations on adjectives after eight months. Gender violations did not yield any effects at any point. For second-year learners (Study 2), a positivity emerged in the posterior region for both subject-verb agreement violations and number violations on adjectives in the first session and remained stable in the second session, with the effect for subject-verb agreement becoming more positive. Similar to Study 1, gender violations did not yield significant effects at any point.

The results show that it is possible to observe effects related to grammatical processing for features shared between the L1 and L2 even after very little exposure to an L2. This result is in line with theories of transfer and is inconsistent with approaches that assume a uniform starting point (N400 or no effects) for the processing of all morphosyntactic structures.
MEMORY RETRIEVAL FAVORS REFERENTIAL ANTECEDENTS: INTERFERENCE EFFECTS IN THE PROCESSING OF REFLEXIVES

Peter Baumann & Masaya Yoshida (Northwestern University).
baumann@u.northwestern.edu

Grammar and working memory both play a critical role in online sentence processing: while some processing phenomena are best accounted for in terms of grammatical constraints (Phillips, 2006), others are due to cue-based memory retrieval (Lewis & Vasishth 2005). However, the exact division of labor between these two processes is still unclear. If online processing is mainly subject to grammatical constraints, structures that do not conform to grammatical constraints will have no influence on online processing. If, on the other hand, online processing is primarily guided by cue-based memory retrieval mechanism, elements that are prominent in terms of retrieval cues but grammatically inaccessible will have an influence on online processing.

In the processing of reflexives, the respective roles of grammatical constraints (command and structural locality) and memory retrieval have been particularly controversial: some studies (e.g. Badecker & Straub, 2002; Patil, Vasishth, & Lewis, 2011) have found slower reading times on the reflexive when there was a gender-matching, but grammatically inaccessible antecedent (1.a.), compared to when there was a gender-mismatching inaccessible antecedent (1.b.). In several other studies (e.g. Sturt, 2003; Phillips, Wagers, & Lau, 2009), no such interference effects have been found.

Critically, in most of the above experiments, there were (at least) two referential antecedents. As a result potential interference effects could be too weak to be observed reliably. However, a reflexive may also be bound by a non-referential element, such as an interrogative pronoun. For example in (2.a./b.), there is a wh-phrase in the embedded clause that contains a reflexive, and a referential NP in the matrix clause. The constraint on reflexive binding (e.g., Binding Condition A: Chomsky 1981) states that in sentences like (2.a./b.), the wh-phrase is a grammatically accessible antecedent, but the referential NP in the matrix clause is not. In this type of configuration, if the search for antecedents is purely structural, it should not be affected by the more referential, but grammatically inaccessible subject NP of the matrix clause. But if the antecedent search is (at least in part) subject to memory retrieval mechanisms, we would expect an interference effect of the referential subject NP.

In a self-paced reading experiment (N=40), we contrasted Clause Type (interrogative clause (2.a./b.) vs. (subject) relative clauses (2.c./d.)), and manipulated the available memory retrieval cues by manipulating the Stereotypical Gender of the matrix subject (mismatch 2.a./c. vs. match 2.b./d.). In relative clauses, the referential head noun and the wh-phrase are obligatorily co-referential, thus, in effect, both the wh-phrase and the head noun are working as the antecedent of the reflexive. As a consequence, the Clause Type manipulation effectively manipulates the referentiality of the reflexive’s antecedent.

We observed that the reflexive and its spill-over were read more slowly when its antecedent is non-referential (main effect of Clause Type). More importantly, we found a significant interaction due to retrieval interference in interrogative clauses only: here the reflexive was read significantly more slowly if the stereotypical gender of the (grammatically inaccessible) subject NP matched the gender of the reflexive than when there was a mismatch.

These results support the position that the processing of reflexives is not solely governed by grammatical constraints, but it also involves cue-based memory retrieval. To reconcile our results with previous findings, we would argue that interference effects are often too small to be observed when both grammatically accessible and inaccessible antecedents are fully referential.

1. a./b. John/Jane thought that Bill owned himself another opportunity.
2. a./b. The janitor/secretary was wondering who wanted to seat herself in the restaurant.
   c./d. The janitor/secretary who wanted to seat herself in the restaurant left soon.
NUMBER ATTRACTION EFFECTS ON OBJECT-CLITIC AGREEMENT IN SPANISH: EVIDENCE FROM NATIVE AND NON-NATIVE SPEAKERS
Adam Zawiszewski (UPV/EHU), Mikel Santesteban (UPV/EHU) & Itziar Laka (UPV/EHU).
adam.zawiszewski@ehu.es

Introduction: Neurophysiological mechanisms underlying subject-verb agreement have been extensively studied in both monolingual [1] and bilingual populations [2]. However, less is known about the electrophysiological responses to number attraction effects (e.g. the key to the cabinets is / *are) during agreement computation by native [4, 5] and particularly non-native speakers. Here we present an ERP study that explored the time course of number attraction errors during object-clitic pronoun agreement processing in native and non-native speakers of Spanish.

Method: In a grammaticality judgment task, 46 native Spanish and 32 highly proficient non-native Spanish bilinguals were presented (word-by-word) with sentences containing a singular object-NP with a local noun that matched (la montaña ‘the mountain’) or mismatched (las montañas ‘the mountains’) in number with the head noun (la casa ‘the house’). Sentences were either grammatical or ungrammatical, depending on whether they contained an object-clitic pronoun that agreed or not in number with the preceding object-NP (e.g., La pastora dijo que la casa de la(s) montaña(s) la / *las visitó en invierno [The shepherdess said that the house in the mountain(s), she visited it / *them in winter]). ERPs (32 electrodes) and grammaticality judgment task responses were registered.

Results: Number attraction effects were found in both behavioral and electrophysiological data. All participants were significantly slower and less accurate responding in mismatch than match conditions (878 ms and 79% vs. 797 ms and 92%). Natives and non-natives showed similar accuracy levels in all number matching conditions (93% vs. 92%). In ungrammatical conditions, nonnatives showed significantly larger attraction effects in accuracy than natives (match-mismatch difference in nonnatives: 16% vs. natives: 25%). Electrophysiological responses of natives and non-natives at the critical word position (la vs. *las clitics; ‘it’ vs. ‘them’) revealed different patterns related to grammaticality effects: Natives showed a fronto-central negativity followed by a P600 in match and only a P600 component in mismatch conditions; whereas non-natives showed only a broad negativity but no P600 in the number match condition, and no effect in the mismatch condition.

Conclusions: Our data reveal similar behavioural and electrophysiological responses for object-clitic pronoun agreement (negativity + P600) than those shown for subject-verb agreement [4] [5]. In native speakers, attraction effects showed to have a deep impact on early (considered automatic) stages of agreement computation (as revealed by the presence of negativity in match and its absence in mismatch conditions), but not on later comprehension processes (P600 unaffected). However, high-proficient non-natives showed to be more prone to attraction errors than native speakers as shown by larger attraction effects in accuracy, and absence of either negativity or P600 components in mismatch conditions. These findings reveal non-natives syntactic processing to be weaker than the one of natives regardless their high L2 proficiency level.

REFERENCES
A COMPARISON OF MUSIC AND LANGUAGE READING: RAPID EYE-MOVEMENT EFFECTS OF MUSICAL INCONGRUENCE
Lauren V. Hadley (University of Edinburgh), Patrick Sturt (University of Edinburgh) & Martin J. Pickering (University of Edinburgh)
l.v.hadley@sms.ed.ac.uk

If people comprehend written music in a similar way to written language, musicians playing from a score should make sense of each note as it is read, processing the notation stream deeply and incrementally. This would mirror the comprehension of written language. One source of evidence for such incrementality in language is the disruption to reading caused by anomalous sentences, for example those involving syntactic anomaly (Ni et al., 1998; Braze et al., 2002) or semantic anomaly (Rayner et al., 2004). Such difficulties emerge in early processing, for example impacting regression rate from the target word, or in time spent reading the target word when first encountered. We therefore report two experiments that investigated musical sight-reading in which pianists played notated melodies which occasionally involved a musical incongruity while their eyes were tracked. If deep processing of notation is incremental, involving the interpretation of harmonic structure before conversion into a motor command, such anomalies should cause rapid disruption to pianists’ eye-movements. If, on the other hand, notation is processed shallowly and is immediately converted into a motor command, then anomalies should show effects either at a later stage or not at all.

We recorded the eye movements of 48 expert pianists reading and concurrently playing novel single-line melodies. A quarter of these melodies included harmonic incongruence, such as the non-resolution of an implied dominant seventh chord. The bar following the implied dominant seventh was designated the target bar. The congruence of the target was counterbalanced across participants, either resolving to the tonic as expected, or being transposed to an incongruent chord (congruence of melodies being confirmed through a pilot study). Notably, harmonic incongruence was subtle, remaining in-key to avoid the visual confound of additional notated accidentals. Significant effects of harmonic anomaly were found both in the rate of regression from the target bar, and in the rate of regression back into the target bar. Pianists looked back from the target bar significantly more often when it was harmonically incongruent, and returned to look back at a harmonically incongruent target bar more often once they had passed it. The former finding indicates a swift detrimental effect of harmonic anomaly, suggesting that deep processing of notation occurs incrementally. However, in contrast to the language literature, we did not find an effect of congruence on time spent in the first-pass of the target bar.

Experiment 1 involved a metronome whereas Experiment 2 did not. The results were broadly similar in the two experiments, both finding significant effects in regression measures, suggesting that the requirement to play with an external beat did not disguise time-based measures of congruence when reading notation. These findings reveal several similarities between reading in music and in language, though there were some striking differences. We additionally found significant pupil dilation for harmonic incongruence in Experiment 1, though not Experiment 2. This is the first time that either regression rate or pupil dilation have been linked to musical congruence during reading, and its relevance in only the metronome experiment indicates that pianists underwent greater processing difficulty when they encountered harmonic incongruence under temporal constraints. Following these findings we suggest that investigating music reading could provide a valuable new perspective from which to view the relationship between music and language processing.
Many previous studies reported that information such as animacy and thematic fit can influence the initial structural analysis (e.g., Trueswell et al., 1993; McRae et al., 1998). In some cases, such an influence is so robust that it completely eliminates the processing cost due to syntactic ambiguity and in other cases it interacts with an influence from fine-grain as well as coarse-grain structural frequency (Trueswell & Tanenhaus, 1994). These results are often taken as evidence for constraint-based processing models, which predict multiple sources of constraints to simultaneously influence the choice of the initial analysis (MacDonald et al., 1994). What is often assumed by these interactive models is the gradience of the effect reflecting thematic fit; the lower the fit of an entity for the thematic role assigned by a dominant analysis, the more activation an alternative analysis would receive (e.g., The suspect detained for questioning was later released). However, it is relatively little known how exactly such thematic information interacts with structural bias in determining the initial analysis before encountering disambiguating information.

To address this issue, we tested relative clause (RC) sentences in Japanese such as (1). The word strings prior to the RC head (yakyusenshu, “baseball player”) were temporarily ambiguous between the main clause (MC) and the RC. The first encountered verb (nondeita, “was drinking”) can only take an animate subject. The sentence-initial noun was either animate and thematically plausible for MC analysis (MC plausible, 1a), animate but implausible (MC implausible, 1b), or inanimate and impossible (MC impossible, 1c). In a visual world eye-tracking experiment, we examined anticipatory looks toward the RC head entity (baseball player) that were launched following the first verb onset but before that of the RC head. The logit of gazes toward the RC head entity relative to the gazes to all the other entities in the scene (i.e., the pictures for nominative NP, RC theme NP, and unrelated distractor) was analyzed using LME models in which Plausibility was included as a fixed effect as well as random slope for the subject and item random variables. If the level of thematic fit with a dominant analysis is inversely correlated with the strength of activation for an alternative analysis, the RC structure would be predicted most in the MC impossible condition and least in the MC plausible condition while the strength of the RC prediction in the MC implausible condition could be somewhere between the two other conditions. On the other hand, it is also possible that the MC analysis is overwhelmingly favored and would be forced even when thematic information strongly biases readers for the alternative RC analysis. This may predict that the RC structure may not be predicted in the MC impossible as much as in the MC plausible condition because the anomalous MC interpretation in the former condition may confuse readers and delay accessing the alternative RC structure.

The results (N = 33) revealed that participants made significantly more anticipatory looks to the upcoming RC head following the onset of the first verb in both the MC implausible and MC impossible conditions compared to the MC plausible condition but significantly less in the MC impossible condition than in the MC implausible. This suggests that an inanimate entity that was thematically impossible agent for the following verb phrase did not trigger an eventually correct RC analysis but caused delay in accessing the alternative. As a measure of real-time processing difficulty, we also analyzed pupillary dilations following the onset of the first verb and found significant increase in pupil size in the MC impossible condition relative to the other two conditions.

Our results together suggest that the level of thematic fit is not linearly correlated with the strength of activation for an alternative analysis; when an utterly anomalous interpretation was forced by strong structural preference, readers could not abandon the incorrect interpretation or access a plausible alternative interpretation.

Examples

1a. Sarariiman-ga biiru-o oisisouni nondeita yakyuusenshu-ni hanasikaketa
   ‘The businessman talked to the baseball player who was drinking beer deliciously.’

1b. Shougakusei-ga biiru-o oisisouni nondeita yakyuusenshu-ni hanasikaketa
   ‘The school kid talked to the baseball player who was drinking beer deliciously.’

1c. Sandaru-ga biiru-o oisisouni nondeita yakyuusenshu-ni watasareta
   ‘The sandals were handed over to the baseball player who was drinking beer deliciously.’
Many studies reported that comprehenders predict upcoming linguistic input using various kinds of information (e.g., Altmann & Kamide, 1999). Recent processing models suggested that estimated probability of upcoming input is inversely correlated with the cost on processing the input; highly predictable input incurs little cost whereas hardly predictable input does large cost (e.g., Levy, 2008). Consistent with the constraint-based lexicalist framework, these models predict that non-syntactic information like pragmatic or lexical semantic information influences the prediction process. The current study examined the prediction of the verb and whether its cost can be influenced by information preceding the verb in the processing of a head-final language (cf. Koniecny, 2000). We conducted two eye-tracking reading experiments with active ditransitive sentences (1a) and passive sentences (1b). The sentences were identical prior to the verb and thus temporarily ambiguous between the active or passive structures. Since the verb in passive form is far less frequent, surprisal-based models predict great processing cost for the passive form verb due to the strong prediction of the active form verb. Our objective is to examine whether the cost of the passive verb could be reduced or eliminated by boosting the probability of the passive structure with information prior to the verb.

In Experiment 1, sentences were either preceded by an adverbial phrase such as *odoroitakotoni* ("surprisingly") or not. The adverbial modifier indicated that the message of the sentence would be somewhat surprising, which may lead comprehenders to prepare for the usually unpredicted passive form verb. This manipulation (Modifier) was crossed with Verb Type. We focus on results from first pass and right-bounded reading time measures in the two regions; preverbal NP region (*aisukoohii-o*, 'iced coffee') for a parafoveal effect and verb region (*hakonda/kakerareta*, 'brought/poured'). Reading times were analyzed using LME models with a backward selection approach. The results for preverbal region showed no effects or interaction. The analysis for verb region showed only a main effect of Verb Type in both measures, revealing greater cost with the passive verb than the active verb. The lack of an interaction, we speculate, was because a sentence can be surprising with an active verb such as *kobosu* ('spill') and the modifier itself did not bias for the passive structure.

In Experiment 2, we replaced the adverbial phrase with an adjective modifying the second argument (*ijiwaruna*, 'mean'). The adjective modifier made the modified entity thematically highly appropriate as agent of adversity events that are typically expressed with Japanese passives. The analysis for preverbal region showed no main effects but revealed a significant interaction in both measures. Note that the interaction can only occur by processing the verb and the effect most likely reflects parafoveal preview of the verb. Further analysis on right-bounded times confirmed the simple effect of Modifier only for passive sentences (mean 95ms) but not for active sentences (21ms), demonstrating that the cost of the unpredicted passive verb was reduced due to the adjective modifier. The analysis for verb region showed only a main effect of Verb Type in both measures, replicating the finding in Experiment 1. This suggests that although the cost on encountering a usually unpredicted verb was reduced by making the passive form verb more expected, the passive structure was still harder to process compared to the active structure.

Our results demonstrated that the cost of the usually unpredicted passive form verb was reduced when information preceding the verb biased readers for the passive structure, as predicted by surprisal-based processing models. The results will be discussed with reference to the estimated probabilities obtained from a sentence completion test.

Examples
1a. (Odoroitakotoni) Kissaten-de ueitoresu-ga (ijiwaruna) okyaku-ni aisukoohii-o hakonda rasii.
   (Surprisingly) at café waitress-NOM (mean) customer-DAT iced coffee-ACC brought
   (active) seems
   "(Surprisingly) it seems that the waitress brought the iced coffee to the (mean) customer at the café."

1b. (Odoroitakotoni) Kissaten-de ueitoresu-ga (ijiwaruna) okyaku-ni aisukoohii-o kakerareta rasii.
   (Surprisingly) at café waitress-NOM (mean) customer iced coffee-ACC poured
   over (passive) seems
   "(Surprisingly) it seems that the waitress poured the iced coffee over by the (mean) customer at the café."
GAMMA FUNCTION MODELLING OF VISUAL WORLD EYE-TRACKING DATA
Emily Atkinson, Akira Omaki, & Colin Wilson (Johns Hopkins University)
atkinson@cogsci.jhu.edu

A large body of research has focused on modeling eye movements during reading. In contrast, relatively few attempts have been made to formally link language processing to patterns of fixation in the visual world paradigm (VWP, cf. [1,2]). Inspired by equations for the BOLD response in the fMRI literature, the current study proposes a model of VWP data in which a time-delayed gamma function is convolved with incoming syntactic information.

In previous work, Mirman et al. have analyzed VWP data with growth curves [3]. However, the orthogonal-polynomial coefficients that result from such an analysis are difficult to interpret in terms of cognitive processes, and the fitted predictions have known artifacts at the asymptotes. Furthermore, growth curve analysis has been applied only to binary data (target vs. competitor), while we are mainly interested in how the activities of multiple possible referents change across processing regions in sentence comprehension. One fully general approach to this issue is provided by the basis function in (1), which is closely related to equations standardly used to model hemodynamic brain responses [4].

\[
p_i(t) = \begin{cases} 
  a \frac{(t-(T_l+T_0))^n}{\lambda^n(n-1)!} e^{-\frac{t-(T_l+T_0)}{\lambda}} & \text{for } t > (T_l + T_0) \\
  0 & \text{for } t \leq (T_l + T_0) 
\end{cases}
\]

where \(a\), \(\lambda\), and \(n\) are parameters that determine the amplitude (\(a\)) and shape (\(\lambda\) and \(n\)) of a scaled gamma function of time (\(t\)). \(T_0\) is the lag between the time of a sentence region (\(T_l\)) and the behavioral response: in the case of BOLD signals, the delay is caused by hemodynamic properties of the brain; here, it represents the time to plan and execute saccades [5]. The value \(p_i(t)\) is the activity (unnormalized probability) of object \(i\) given these parameters. In the model, there are multiple instances of (1) that collectively modulate the activity of referents in the VWP display across several sentence regions. Predictions of fixation proportions are derived by normalizing \(p_i(t)\) across referents within times.

The data to be modeled comes from a VWP study on filler-gap dependency processing [6] in which the target question followed a story about the displayed images [7]. Visual stimuli consisted of five referent images; fixations were recorded while participants listened to temporarily ambiguous \(wh\)-questions (Can you tell me what / Emily\(_{R1}\) / was | eating\(_{R2}\) / the cake\(_{R3}\) / with\(_{R4}\) ?). The referents in a display were the subject, the target object and instrument, and a distractor object and instrument. As the distractors are never mentioned, the critical fits are for fixations to the subject, target object (cake), and target instrument (fork, which corresponds to the filler). Four syntactic regions are used to predict fixations: the subject (R1) and the object (R3) bias fixations on their respective referents, the verb (R2) on the object, and the preposition (R4) on the instrument. The parameters of region-specific instances of (1) were fit to human fixation data (Fig.1, averaged across trials and participants). The resulting fit (Fig.2) accounted for the main qualitative patterns observed in the data and was superior to a baseline uniform distribution both overall (log likelihood: -68105.7 vs. -112412.3) and for the critical items (-30101.85 vs. -47461.56).

Figure 1. VWP data
Figure 2. Model predictions

GRAMMATICAL FAITHFULNESS IN PROCESSING OF MULTIPLE DEPENDENCIES
Emily Atkinson & Akira Omaki (Johns Hopkins University)
atkinson@cogsci.jhu.edu

Parsing operations must be confined within the limits of working memory resources; thus, large memory demands associated with multiple incomplete dependencies cause a ‘processing breakdown’ [1], but the nature of the breakdown is poorly understood. For example, doubly-nested sentences that omit one of the required verbs are ungrammatical, but are perceived as grammatical [2]. This grammatical illusion could indicate that when processing multiple incomplete dependencies, the parser may be grammatically unfaithful, resorting to an ungrammatical representation in which a filler lacks thematic interpretation. Alternatively, the parser may become unfaithful to the input [3,4] by ignoring or deleting the entire dependency, which reduces the memory load but maintains a grammatical representation. The current study uses sentences that offer a potential ‘short-cut’ for reducing the memory demand to test whether processing multiple dependencies compromises grammatical faithfulness. Crucially, this short-cut would result in a violation of the island constraint, which the parser generally respects [5,6]. Results from two experiments suggest that the parser adheres to the grammar even under large memory load.

Exp1 (self-paced reading, N=34) used sentences with three long-distance dependencies (1): a fronted adjunct clause, a backward anaphora, and a filler-gap dependency. The presence of three dependencies resembles the memory load in doubly-nested sentences, but (1) allows an early integration of the filler if the island constraint is violated. We manipulated the semantic fit between the filler and the verb, the presence of an island, and pronoun gender. The gender mismatch manipulation (he/she...Mike [7]) is used to verify if the antecedent expectation is maintained. The critical manipulations are the crossing of filler-plausibility and islandhood. It is widely observed that the parser creates a gap at the first permissible gap location (active gap-filling: [5,6]); thus, for non-island conditions we expect a plausibility mismatch effect, i.e., a slowdown on or after the verb in the implausible-filler conditions (1ef:which city...wrote) compared to the plausible-filler conditions (1ab:which book...wrote). If the parser respects island constraints, the plausibility mismatch effect should disappear in the island conditions (1cd vs. 1gh). If the grammatically unfaithful strategy is used, a plausibility mismatch effect is expected regardless of islandhood.

The gender mismatch effect appears immediately after the main clause subject [F=7.33,p=.01], confirming that the antecedent expectation was maintained. The critical verb region showed longer reading times for non-island conditions [F=4.74,p=.04] but no significant interaction [F<1]. The spillover region showed main effects of island & plausibility [Fs=8.79,ps<.007], and crucially, a significant interaction [F=5.91,p=.02]. Pairwise comparisons confirmed the presence of a plausibility mismatch effect in non-island conditions [t(33)=-3.14,p=.004] and its absence in island conditions [t(33)=-1.24,p>.1].

Exp2 (N=12) tested whether the island effect robustly holds in an offline 7-point-scale acceptability judgment task. (1c) was contrasted against a version that includes a preposition (e.g., about) within the island forcing a violation. The grammatical version is significantly more acceptable than the ungrammatical one [5.02 vs. 3.65; t(11)=3.33,p=.007], suggesting that the island constraint is respected regardless of the task modality.

These results indicate that during processing of multiple incomplete dependencies, the structure building mechanisms remain faithful to grammatical constraints but not necessarily to information in the input. Implications for the relation between syntactic processing and perceptual noise [3,4] will be discussed.

References:
Similarity-based interference has been identified as an important determinant of difficulty in resolving linguistic dependencies [1]. This difficulty is characteristic of cue-based retrieval — the retrieval cues cannot unambiguously diagnose the target encoding. For instance, [2] showed that processing a verb is difficult when its subject is a complex constituent containing other subject phrases. In our study, two reading experiments and a computational simulation were designed to test the role of case/clause-finiteness in subject retrieval, two features correlated with subjecthood in English. The RT results indicate that case/finiteness, not structural position, are the operative cues. Combined with our simulation, they suggest that the retrieval of clauses, not just subjects, is the cause of interference in resolving a subj-verb dependency.

Simulation.— We simulated the attachment of complex subjects in ACT-R, deriving the retrieval schedule using a left-corner parsing algorithm [3]. In our simulation, encountering the subject triggered the creation of a clausal node encoding the expectation for an upcoming verb. At the verb, the clause was retrieved and linked with the input. We found that it was actually possible to retrieve the correct nominal, due to an activation boost in early processing. In contrast, retrieving the correct clause suffered interference from intervening clauses [2, cf. 4]. This suggests that previous findings of difficulty from intervening subjects stem from similarity between matrix and embedded clauses, not their subjects per se.

Design.— In self-paced reading (n=40) and eye-tracking (n=30), we tested the role of clause-type/case information in subject retrieval. A 4 × 2 design compared: (i–ii) (un)ambiguous finite clause-complements with nominative subjects; (iii) non-finite ECM constructions with accusative subjects; and (iv) non-finite object-control constructions with an implicit subject. The four constructions were fully crossed with distractor subject size: we reasoned that larger subjects would accrue more activation and interfere more strongly at the matrix Aux+V (underlined).

Results.— In eye-tracking go-past times, there was a V-Type × size interaction: large distractor subject conditions showed a reliable slowdown, except for non-finite, ECM conditions. In total times, only the conditions containing finite intervening clauses (i–ii) were read more slowly than conditions with non-finite intervening clauses (iii–iv). The go-past pattern was also replicated in the self-paced reading times.

Discussion.— Our findings support a model in which finiteness (or case) plays a primary role in S-V dependency resolution. The absence of interference from ECM distractors undercuts a role for structural position alone. The interaction of V-type and Size supports a notion of representational strength, like activation: the additional processing required for a large distractor subject boosts its likelihood of retrieval when it matches the cues. The reading patterns indicate that the retrieval cues include finiteness (for clauses) and/or case (for subjects); the simulation adds weight to the first possibility. Also supporting clausal retrieval is recent evidence from agreement-attraction that case plays a role only when it is morphologically visible [5], unlike in our study. Requiring further research is the fact that the object control conditions did give rise to a size effect, though not in total times: one possibility is that, in that construction alone, the distractor NP was linked to two roles in the argument structure.

Eye Movements During the Processing of Pronominal Relative Clauses
Douglas Roland (The University of Tokyo), Yuki Hirose (The University of Tokyo),
Gail Mauner (The University at Buffalo), & Stephani Foraker (Buffalo State College)
doug.roland@gmail.com

The processing of pronominal relative clauses plays a key role in theories of sentence processing. Object relative clauses such as (1) are typically read more slowly than subject relative clauses such as (2), when the embedded noun phrase is a full NP, such as the pilot. However, when the embedded noun phrase is a pronoun, such as I or me, object relatives such as (3) are read more quickly than the analogous subject relative (4). This change in processing difficulty has been attributed to both frequency and discourse expectations (e.g., Reali and Christiansen, 2007; Roland et al., 2012), and has been used as an argument against memory based accounts such as Warren and Gibson (2002) and Gordon et al. (2001), because memory accounts, while predicting that pronominal object relatives should be less difficult than full NP object relatives, cannot explain why pronominal object relatives would be easier than pronominal subject relatives. We investigated the effects of relative clause (subject vs. object) and relative clause noun phrase (pronominal vs. full) on eye movement patterns for 61 native speakers (L1), and 24 Japanese (L2) speakers of English.

**Predictions:** For L1 speakers reading full NP relative clauses, we would expect to replicate Staub (2010), who found an increase in regressions at the object relative clause noun phrase, which he attributed to a violation of expectations, and an increase in fixation lengths at the object relative clause verb, which he attributed to memory load effects. For pronominal relative clauses, discourse/frequency accounts would predict differences primarily at the pronoun, while memory based accounts would predict longer reading times only at the embedded verb of full NP object relative clauses. Baek, Ionin, and Kim (2011) found that, like L1 speakers, Korean L2 speakers of English had longer reading times for object relative clauses than for subject relative clauses when the clauses modified animate nouns, but not when they modified inanimate nouns, suggesting that the L2 speakers were sensitive to the same frequency biases as L1 speakers. If the L2 speakers have learned that subject relatives are more frequent than object relatives, but have not yet learned the finer-grained distinction based on NP type, we would expect to find that eye movements for both pronominal and full NP relatives are similar to those for L1 full NP relative clauses.

**Results:** For the L1 full NP relative clauses, we replicated the overall patterns of eye movements found by Staub (2010), suggesting that both memory and expectations play a role in the processing of full NP relative clauses. For L1 pronominal relative clauses, we found that the object relative clause pronouns were read more quickly than subject relative clause pronouns, and no differences in eye movements for the relative clause verb. Thus, the L1 results suggest that processing difficulties at the relative clause noun phrase are primarily expectation-based, while processing difficulties at the relative clause verb, found in full NP object relative clauses, are consistent with memory-based accounts.

For the L2 speakers, we found that the pronoun in pronominal object relatives was read more quickly than the pronoun in subject relatives, but that there were no significant differences in eye-movements at the relative clause verb (i.e., we found no evidence that full NP object relatives were more difficult than full NP subject relatives for the L2 speakers).

Overall, our results suggest that the reversal in processing bias for pronominal relative clauses is primarily due to differences in the processing of the pronouns themselves, and that the reversal in difficulty does not show up on relative clause verb (unlike in self-paced reading, where both the embedded verb and pronoun are faster for object relatives). That we see facilitation for pronominal object relatives within the same L2 data for which we do not see significant processing difficulties for full NP object relatives suggests that the bias for pronominal object relatives is stronger than the bias against full NP object relatives.

1. The flight attendant that the pilot recognized forgot to serve drinks.
2. The flight attendant that recognized the pilot forgot to serve drinks.
3. The flight attendant that I recognized forgot to serve drinks.
4. The flight attendant that recognized me forgot to serve drinks.
The Expectation Effect of Roles and Role Fillers in Korean
Hongoak Yun (Konkuk University), Dongsu Lee (Konkuk University), Yunju Nam (Konkuk University), Hye-in Jeong (Konkuk University), Eunyoung Chung (Konkuk University), Upyong Hong (Konkuk University).
oaktreepark@gmail.com

Does the expectation of which thematic role (role expectation) occurs have an effect on sentence comprehension independently of the expectation of which particular word (role filler expectation) is likely to appear? Considering previous findings that case markers in head-final languages play an anticipatory role to elicit role expectation (Kamide et al., 2003), we investigated whether the expectation of roles had an additional effect on sentence comprehension above and beyond the expectation of particular role fillers, using Korean.

We conducted 2 Role Expectation (High, Low) x 2 Role Filler Expectation (High, Low) self-paced reading with a judgment and ERP studies. The expectation degree of roles and role fillers was determined through cloze tasks: Patients with accusative case markers were strongly expected to occur after recipients with dative case markers appeared (about 70%), whereas recipients were weakly expected to appear after patients occurred (about 5%). In each role condition, role fillers were either highly likely or unlikely. We hypothesized that processing difficulty of role fillers would be determined not only by how likely they were in a given context but also by how strongly roles associated with role fillers were expected.

<table>
<thead>
<tr>
<th>Role Filler</th>
<th>R1</th>
<th>R2</th>
<th>R3 (Target)</th>
<th>R4</th>
<th>R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) High High Chelswu-ka Chelswu-NOM</td>
<td>wuncensa-eykey</td>
<td>driver-DAT</td>
<td>chapi-lul</td>
<td>fare-ACC</td>
<td>emchengnakey</td>
</tr>
<tr>
<td>(2) High Low Chelswu-ka Chelswu-NOM</td>
<td>oyyaswu-eykey</td>
<td>player-DAT</td>
<td>chapi-lul</td>
<td>fare-ACC</td>
<td>emchengnakey</td>
</tr>
<tr>
<td>(3) Low High Chelswu-ka</td>
<td>chapi-lul</td>
<td>fare-ACC</td>
<td>wuncensa-eykey</td>
<td>driver-DAT</td>
<td>emchengnakey</td>
</tr>
<tr>
<td>(4) Low Low Chelswu-ka</td>
<td>chapi-lul</td>
<td>fare-ACC</td>
<td>oyyaswu-eykey</td>
<td>player-DAT</td>
<td>emchengnakey</td>
</tr>
</tbody>
</table>

As hypothesized, unlikely role fillers whose roles were weakly expected were the most difficult to read, whereas likely role fillers whose roles were strongly expected were the easiest to read. Role expectation (β = 8.95, S.E. = 4.02, t = -2.23) was effective in addition to role filler expectation (β =-138.48, S.E. = 41.31, t =- 3.35), while the effects of lexical frequency and length were controlled for. Of great interest, we observed N400 effect in left medial areas, only when roles were strongly expected (t (32) = 2.5, p < .05). The amplitudes for unlikely role fillers were more negative than those for likely role fillers. However, when role expectation was weak, P600 effect emerged in left medial areas. The amplitudes of unlikely role fillers went more positive than those of likely role fillers (t (32) = -2.2, p < .05).

Our asymmetric ERP patterns suggest that the nature of processing difficulty for role fillers differs depending on when the slots for yet-to-encounter roles associated with role fillers are introduced and how strongly particular role fillers are expected in a given context. Upon the recognition of recipients in (1 & 2), slots for patients are introduced, prior to the occurrence of targets. Then, the processing difficulty of patient fillers is dominantly affected by the degree of semantic expectation of the fillers. However, in encountering recipients after patients in (3 & 4) (i.e., slots for roles are not anticipatorily introduced), readers seem to have structural difficulties in integrating those unexpected roles into sentences, and they feel easier to integrate semantically likely role fillers than unlikely role fillers. We reason that the expectation of roles triggered by linguistic information and that of role fillers by the contextual information are both important in sentence comprehension (c.f., Kuperberg, 2007).


Interference in long-distance dependencies is core to Relativized Minimality (RM, Rizzi, 2001) and cue-based (CB) memory retrieval models (e.g., McElree et al., 2006). According to RM, a local relation between an extracted element and its trace is disrupted by an intervening element having the same syntactic features of the extracted element. According to CB, retrieval interference is generated under the same conditions, although both syntactic and semantic features can engender interference. Two acceptability studies on wh-islands in English and French (Atkinson et al., 2013; Villata et al., 2013) showed that sentences with two lexically restricted wh-elements (2) are more acceptable than those containing two bare wh-elements (1), despite that both have identical sets of syntactic features ([+Q, +N] in (2) and [+N] in (1)). Although this finding runs against RM, it is in line with CB, which predicts semantic distinctiveness of lexically restricted wh-elements to engender less retrieval interference than bare wh-elements, which will lead to higher acceptability rates. Here, we used the speed-accuracy trade-off (SAT) procedure to investigate how interference affects memory retrieval in resolving long-distance dependencies by tracking response accuracy across the full time-course of retrieval. Our main contrasts fully crossed lexically restricted and bare wh-elements in structures in which a wh-element (Wh1) has been extracted over a second wh-element (Wh2), as illustrated in conditions (1)-(4):

1. Bare Identity: What do you wonder who built __?  
2. Complex Identity: Which building, do you wonder which engineer built __?  
3. Inclusion: Which building, do you wonder who built __?  
4. Inverse Inclusion: What, do you wonder which engineer built __?  

Thirty-six sets of contrasts like (1)-(4) were generated, along with 4 additional contrasts consisting of corresponding sentences without extraction (e.g., Who/Which wonders who/which engineer built this building?). For each of the 8 sentence types, we generated a corresponding ungrammatical variant (e.g., What do you wonder who slept?). Sentences were presented one phrase at the time, and participants were asked to make yes/no acceptability judgments at each of 18 tones presented at 250ms intervals after the onset of the last phrase. Figure 1 shows the average full time-course functions for acceptability judgments of the 15 participants tested for conditions (1)-(4) in d' units (hit rates for (1)-(4) scaled against false alarm rates for ungrammatical variants). An ANOVA on asymptotic values shows significant main effects of both Wh1 (F(1,14)=31.74, p<.001) and Wh2 (F(1,14)=20.83, p<.001), with higher d' values for restricted wh-elements, and no interaction. These asymptotes' patterns conform to the prediction of CB that lexical restriction serves to make constituents more distinctive in memory. Interestingly, analysis of processing speed (SAT dynamics) revealed that Bare Identity was processed, on average, twice as fast as the other 3 conditions, despite the low asymptote. Inspection of individual subjects’ functions for this condition (Fig. 2) revealed that this difference is due to a clear non-monotonicity in the dynamics, absent in the other conditions: 11 of the 15 subjects showed disproportionally high acceptance rates for Bare Identity early in processing, which were reversed later in processing, suggesting late interpretative problems. Although it is possible that syntactic constraints of the type proposed in RM are operative late in comprehension, it is more plausible to see the late rejection of Bare Identity as due to the difficulty of constructing an informative interpretation (Abrusan, 2014).
NAVIGATING INTERFERENCE EFFECTS IN AGREEMENT IN SENTENCE COMPREHENSION

Julie Franck (University of Geneva), Maria Lucia Cali (University of Geneva) & Luigi Rizzi (University of Siena and University of Geneva)

julie.franck@unige.ch

Empirical research on object interference in agreement production supports the view that features spread through the hierarchical structure under the guidance of syntactic constraints like intervention and c-command [1]. In these studies, a distracting plural feature mismatching the agreement controller has a detrimental effect on agreement computation, causing ‘attraction’. In contrast, recent studies in sentence comprehension suggest that a mismatch in agreement features may actually produce a facilitatory effect [2], in line with similarity-based interference effects in sentence comprehension [3,4]. However, the production vs. comprehension contrast fails to capture the whole picture, since the comprehension literature also reports attraction-like, detrimental effects of feature mismatching [5,6]. We report three experiments on object interference in sentence comprehension in French, which suggest that the apparent contradictions in the comprehension literature result of the variability in the tasks used, tapping into different grammatical processes. Whereas self-paced reading taps into cue-based memory retrieval at play in structure building, which is enhanced by featural distinctiveness, grammaticality judgment and two-choice response time paradigms tap into feature checking which is fooled by the presence of distracting features. Moreover, we show that the process of feature checking is sensitive to the same syntactic factors that affect agreement production.

Experiments 1 (self-paced reading) and 2 (grammaticality judgment) contrasted interference from the plural moved object of the embedded clause ‘patients’ in the agreement of the embedded verb ‘cures’ in object relatives (e.g., John speaks to the patient(s), that the medicine cures _i_) to interference from the unmoved object of the matrix clause in superficially similar sentence complements (e.g., John says to the patient(s) that the medicine cures). Linear mixed-effects models on response times show significant interference in object relatives but not in sentence complements. However, whereas the plural ‘patients’ had a facilitating effect in Experiment 1, it had a reversed, detrimental effect in Experiment 2. Taken together with the additional finding that interference interacted with animacy features, data from Experiment 1 suggest that the self-paced reading task taps into a cue-based retrieval process in which number and animacy are used as cues for structure building. In contrast, the reverted, attraction effect in Experiment 2 replicates findings from sentence production with the same structures [1]. This suggests that the grammaticality judgment task taps into the process of agreement feature checking itself, and that this process shows the hallmarks of structure-sensitivity found in production. Experiment 3 (grammaticality judgment) further explores structure-sensitivity by manipulating object relative clauses with complex object NPs involving a prepositional phrase modifier (e.g., Which patient(s) of the doctor(s) do you say that the lawyer defends?). Results show attraction from the plural feature on the highest, c-commanding NP (‘patients’) but not from the lowest, preceding modifier (‘doctors’), again attesting to the structure-dependent nature of interference in agreement comprehension.

We conclude that (i) the direction of the interference effect (similarity-based vs. attraction) in sentence comprehension depends on the task used and the underlying process that it presumably taps into, and (ii) feature-checking in agreement in sentence comprehension is structure-dependent.

SMALL CLAUSE AVAILABILITY TRIGGERS HIGH ATTACHMENT IN ENGLISH

RC Attachment. RC Attachment. Preferences for Relative Clause (RC) attachment ambiguity resolution differ across languages: Low Attachment (LA) in English and High Attachment (HA) in Spanish (Cuetos & Mitchell, 1988). Grillo and Costa (in press) (G&C) show that previous work on RC attachment overlooked the ambiguity between an RC and a Pseudo Relative interpretation (PR, see Cinque 1992) in HA languages (1a,2a).

1 a. PR interpretation 1 b. RC interpretation
Sento [PR [l ragazzo [che piange]]]
I heard [SC [the boy [crying]]]
= I heard an event of the boy crying
Sento [DP il [NP ragazzo [SC che piange]]]
I hear [DP the [boy [SC that is crying]]]
= I hear the unique boy that is crying

(2) a. Sento [PR il figlio; dell'uomo che parla] b. I hear [SC the son of the man talking]

PRs and RCs are string identical, but differ in structure and interpretation distinct (1a,b). PRs correspond to Small Clauses (SC) in English (2b): both are eventive, selected by perceptual predicates and only compatible with HA interpretations (indexes in 2). Stative predicates (eg, live with), on the other hand, only select for RCs. G&C propose that variation in RC attachment is due to the asymmetric availability of PRs crosslinguistically, and as PRs-predicates are easier to parse than RC-modifiers, a PR interpretation (ie, HA) is preferred (all else being equal). Two previous experiments in Italian showed HA in PR/RC contexts and LA in RC-only contexts.

Experiment 1: Syntax or Predicate Pragmatics? The current data cannot rule out the possibility that the PR-effect is pragmatically driven by the predicate-type manipulation rather than a syntactic preference, as claimed by G&C. To test this alternative account, the same predicate manipulation in a nonPR language, like English, is necessary. LA across all predicates would be consistent with the G&C account. Experiment 2: Can English become a HA Language? G&C predict HA of reduced RCs in English (string identical to SC) with perceptual predicates and LA with stative predicates. General Methods. In two English timed questionnaires (N=60) we contrasted SC-compatibility (perceptual vs. stative predicates). Experiment 1 tested full RCs while Experiment 2 contained a reduced version of those same RCs (Table 1). The manipulation was crossed with a manipulation of Environment (Nominal vs. Verbal). As environment had no effect, we have excluded from the Table of results.

<table>
<thead>
<tr>
<th>Experiment 1</th>
<th>%HA</th>
<th>Experiment 2</th>
<th>%HA</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC only Mary heard the son of the man that was talking.</td>
<td>32.5</td>
<td>SC/RC Mary heard the son of the man talking.</td>
<td>56.1</td>
</tr>
<tr>
<td>RC only Mary lives with the son of the man that was talking.</td>
<td>37.2</td>
<td>RC only Mary lives with the son of the man talking.</td>
<td>21.1</td>
</tr>
</tbody>
</table>

Table 1

Attachment Preferences Results: As predicted, we obtained across the board LA in Exp. 1 and HA in SC-compatible conditions in Exp. 2 (Table 1). We conducted separate mixed effects logistic regression analyses for each experiment. SC-compatibility and Environment were fit as fixed factors, and subject and item as random factors. Despite the overall LA preference in Exp. 1, the effect of SC-compatibility was significant in both Experiments (p<.001; p<.0001).

RT Results: As predicted, RTs varied with attachment: in Exp. 1 we found significantly longer RTs for HA than for LA across all conditions (p=.0001). In Exp. 2 we observed a significant interaction between SC-compatibility and Attachment, with significantly longer RTs for HA limited to the RC-only conditions. Both effects of Attachment and RT strongly interact with Experiment and SC-compatibility (p<.0001).

Discussion: These data demonstrate, for the first time, that HA preference emerges even in "LA languages" like English once an SC analysis is made available. SC-compatibility also increased frequency of HA for unambiguous RCs, without overcoming the overall LA preference. While pragmatics (and/or local SC-compatibility) plays a role, the combined results fully support a syntactic account of the effects.
COMPREHENSION PRIMING EVIDENCE FOR ELLIPTICAL STRUCTURES
Julian Grove (University of Chicago), Emily Hanink (University of Chicago), & Ming Xiang (University of Chicago)
juliang@uchicago.edu

It is debated whether or not syntactic structure is necessary for the interpretation of verb phrase ellipsis (VPE) (Merchant, 2009), as various theoretical accounts of ellipsis posit covert structure (e.g., Merchant, 2001; Fiengo and May, 1994), or null anaphora (Hardt, 1993). We present preliminary results from a comprehension-priming paradigm to suggest that processing VPE indeed activates syntactic structures.

A self-paced reading paradigm was implemented in which a Prime sentence was presented as a separate trial immediately preceding a Target trial. The Target always contained a verb that had the option to continue either into a sentential complement (SC) or a direct object (DO) structure. The verbs were chosen (and normed) to be strongly DO-biased, but the target sentence was ultimately a SC structure (see Table 1). The Prime sentence consisted of two clauses, instantiating a 2x4 factorial design. The first factor is the structure of clause 1 (two levels): it either contained the same SC structure as the Target, or contained only a direct object structure; in either case the Prime and the Target share the same verb. The second factor was the structure of clause 2 (four levels): it varied between (i) VPE; (ii) ‘do-it’ anaphora; (iii) a full clause replicating the clause 1 structure; or (iv) an intransitive predicate.

As the matrix verb in the target sentence (e.g., “proclaimed” in Table 1) is DO biased, a DO structure prime will further strengthen that DO bias. On the other hand, an SC Prime will shift the probabilistic verb bias towards a SC or neutral bias. Previous literature has long established that verb bias information is used immediately to guide online ambiguity resolution (e.g., Garmsey et al., 1997). We predict that RTs on the post-verbal NP, which is ambiguous between a direct object and a sentential subject parse, will be shorter when the DO bias is made even stronger by a DO prime; RTs will be longer when the original DO bias is weakened by a SC prime, leading to more uncertainty on the NP. In the mean time, we also expect an interaction between the type of Clause 2 and the priming effect on the post-verbal NP: if syntactic structures of the antecedent are activated in interpreting VPE, we expect a similar priming effect from VPE and full-clause primes. If VPE is resolved like anaphora (e.g., semantically), VPE and ‘do-it’ anaphora should pattern alike.

Table 1: Stimuli Example (items n = 40; subjects n = 63)

<table>
<thead>
<tr>
<th>Clause 1</th>
<th>Clause 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prime: The students planned to proclaim the funding news was false. and later that year they did.</td>
<td>V NP Aux Pred and later that year they did it.</td>
</tr>
<tr>
<td>The students planned to proclaim the funding news. and later that year they did proclaim the funding news was false.</td>
<td>V NP and later that year they were afraid.</td>
</tr>
<tr>
<td>Target: The news reporter proclaimed the royal birth was official during the broadcast.</td>
<td>V NP Aux Pred</td>
</tr>
</tbody>
</table>

A mixed effect model (n=63) showed that on the target sentence, the critical NP region exhibited a marginal interaction between the two experimental factors (p<0.09). Within the targets after VPE primes, the NP was read slower when primed with a SC vs. a DO (log-transformed RTs, β=0.059, se=0.026, |t|>2). We found no priming effect for the ‘do-it’ primes on this region. The contrast between VPE and ‘do-it’ anaphora provides evidence suggesting that the immediate priming effect on the ambiguous NP from VPE primes is probably due to the rapid probabilistic adaptation from the reactivated syntactic representation at the ellipsis site; whereas, the semantically resolved ‘do-it’ anaphora Primes do not have the same impact on the Target sentence.

Somewhat surprisingly, in contrast to the VPE Primes, comparable immediate priming effects were not obtained for the full-structure Prime. But a sentence-rating task conducted on MTurk on the same set of Prime stimuli (subj n = 40) revealed a sharp acceptability contrast between items containing full structures in the second clause, and those containing ellipsis (on a 1-7 scale, for Sentential-Complement Primes, full structure: 3.87; ellipsis: 4.81; For DO Primes, full structure: 4.71; ellipsis: 5.57; all ps < 0.01). It is possible that the redundancy and unnaturalness of the full-structure Prime hindered the hypothesized priming effect on the Target. We are conducting follow up studies to examine this possibility.

In summary, the present data constitute the first piece of reported evidence that sentences containing VPE not only facilitate structural priming in comprehension (for production priming evidence on VPE, see Xiang, Grove, and Merchant, 2013; and Cai, Pickering and Sturt, 2012), but do so to a greater degree than sentences containing ‘do-it’ anaphora, indicating the role of syntactic representations in the interpretation of VPE.
ADDING A THIRD WH-PHRASE SOMETIMES INCREASES THE ACCEPTABILITY OF MULTIPLE-WH-QUESTIONS
Lauren Ackerman & Masaya Yoshida (Northwestern U)

**Introduction:** There is an appreciable difference in acceptability between some multiple wh-questions with two and three wh-words. The amnestying effect of the third wh-word (the so-called "additional wh-effect") is noted in much of the syntax literature [1,2] in the context of the violation of the Superiority Condition ([3]), which states that lower wh-phrases cannot move over higher wh-phrases. The observation in the syntax literature has been that an additional third wh-phrase can ameliorate the Superiority Condition violation. The psycholinguistics literature, however, does not support this observation in certain environments [4,5]. This study asks whether the additional third wh-phrase ever improves acceptability of otherwise unacceptable multiple wh-questions, and if so, in which environments. Such a study, we believe, can help reveal the nature of the constraints that are operative during acceptability judgments, and the reliability of the acceptability judgments reported in the literature.

**Background:** [5] finds that a third wh-phrase does not increase the acceptability of object-initial wh-questions, as in (1).

(1) a/b. Peter and his friends were trying to remember what who carried { ø / when }.

[5] reports that the higher acceptability predicted by the additional wh-effect is not observed through formal acceptability judgment studies. Besides examples like (1), other patterns have also been recognized in the literature. For example [2,6] report the additional wh-effect in (2), which, as far as we know, has not been experimentally investigated.

(2) a/b. You know perfectly well where who { ?put what/ *put it }.

In (2), the superiority violation is caused by moving the lower complement wh-phrase where over the subject wh-phrase who. Thus (1) and (2) differ in what causes the superiority violation and in verbal argument structure. Although some judge (2a) to be better than (2b), the difference is not clearcut. We test (2) using formal experiments to check whether or not a third wh-phrase ever ameliorates superiority violations. To check whether the argument structure of the verb affects the additional wh-effect, we compare (2) against (3), which uses montransitive verbs, thus the additional wh-phrase is a temporal adjunct like (1), but maintains the movement of the locative complement wh-phrase where over the subject wh.

(3) a/b. You know perfectly well where who { went when/ went ø }.

If the additional wh-effect is not real, we expect (2/3a) should not be judged better than (2/3b), but if the effect is real, we expect (2/3a) to be judged better than (2/3b). Moreover, if the additional wh-effect is real and if the argument structure is a crucial independent factor, we expect a different pattern between (2) and (3).

**Experiment** (N=80): We conducted a force-choice task on Amazon Mechanical Turk, manipulating Argument Structure (Ditransitive vs Montransitive) in a 1x2 design. Subjects were instructed to choose the more acceptable option from a minimal pair of sentences.

**Results:** The results were analyzed using a generalized linear mixed model fit by the Laplace approximation. We find a significant effect of Argument Structure (β=−0.87, SE=0.11, z=−7.93, p<0.0001). This indicates that the two conditions behave differently regarding their preference for an additional wh-word. A 1-sample proportions test with continuity correction revealed that the third wh-word option was chosen significantly less than chance for montransitive conditions, at 41.8% (χ²(1)=28.22, p<0.0001), and significantly more than chance for ditransitive conditions, at 56.8% (χ²(1)=16.34, p<0.0001). This indicates that ditransitive verbs display a subtle but significant preference for the additional wh-word, whereas montransitive verbs prefer only two wh-words.

**Conclusions:** The presence of the additional wh-effect in ditransitive but not montransitive verbs suggests that the argument structure of the verb plays a role in the distributions of the additional wh-effect. This indicates that the additional wh-effect and something like the Superiority Condition are operative. While this observation is progress toward understanding in what environments the additional wh-effect can occur, future work must attempt to establish the precise role of the verbal argument structure in multiple wh-questions.

REPEATED EXPOSURE TO A WORD MEANING PROTECTS AGAINST INTERFERENCE FROM EXPOSURE TO ITS ALTERNATIVE MEANING

Hannah Betts (University College London), Garry Cai (University College London) & Jennifer Rodd (University College London).
hannah.betts.10@ucl.ac.uk

Presenting an ambiguous word (e.g. ‘bark’) in a context that biases disambiguation towards a particular meaning (e.g. ‘tree bark’) can bias how the meaning of the ambiguous word is interpreted 20 minutes later [1]. This meaning priming effect suggests that the mechanism underlying meaning preferences is flexible based on recent experience. The present experiment explored the nature of the underlying learning mechanism by investigating the impact of three factors: whether meaning priming reduces over time, whether meaning priming reduces with subsequent exposure to language and whether these factors are affected by the number of exposures to the ambiguous word meaning.

Participants heard homophones disambiguated towards their less frequent (subordinate) meanings (e.g. ‘the woodpecker clung onto the bark’). Homophones were presented once (low priming) or three times (high priming). Then, participants encountered a filler task involving either no language, language unrelated to the homophones, or the homophones in the context of their more frequent (dominant) meanings (e.g. ‘the cat jumped at the sound of the bark’). The filler task created a 30-min delay (compared to a no filler control, 15-min delay) between priming and the subsequent word association test, in which meaning preferences for the homophones were measured.

Logistic mixed effects modelling of the data revealed a main effect of priming: compared to an unprimed baseline, the two (low and high) priming conditions elicited more responses related to the subordinate meaning. The priming effects were comparable for the two priming conditions after the 15-min no filler delay, and after the 30-min delay with no language or unrelated language filler tasks. The meaning priming effect, however, was significantly reduced after a 30-min delay with the dominant meaning filler task, and low priming was more vulnerable to the interference of the dominant meaning than high priming.

It seems that the impact of encountering the subordinate meaning of a homophone does not significantly weaken within half an hour or with subsequent exposure to unrelated language. However, an encounter of the dominant meaning weakens, and even eliminates, the impact of an earlier encounter with the subordinate meaning. These findings can be accommodated by the strength of connections between the ambiguous word form and its multiple meanings proposed in the connectionist model [2]. The model, however, is insufficient in accounting for the finding that whilst three encounters of the subordinate meaning did not increase meaning priming compared to one encounter, it does seem to provide resilience against the interference of the dominant meaning encounter. Such a finding requires a resilience dimension in the meaning-form link, which strengthens as a function of exposure to a meaning and protects against interference from an alternative meaning. While it is currently unclear how such a dimension can be implemented computationally, the present findings suggest that more than one mechanism may underlie meaning priming of ambiguous words.

DISFLUENCY AND ATTENTION IN LANGUAGE COMPREHENSION.
Sam Miller, Martin Corley, Martin Pickering, Ian Finlayson and Eleanor Drake (University of Edinburgh).
S.J.Miller-6@sms.ed.ac.uk

Spoken language comprehension is impacted by the presence of disfluencies. Listeners are more likely to remember words that follow disfluency (Corley et al., 2007). The presence of disfluency affects the attentional state of the listener, as indexed by attenuation of event related potentials to acoustically manipulated words post disfluency (Collard et al., 2008). It has therefore been suggested that low-level attentional effects of disfluencies may underlie their positive impact on language comprehension (Collard et al., 2008). This attentional account states that upon encountering disfluency, language users employ more attentional resources to help them resolve the situation. In the current study, we ask how disfluencies impact listener attention at a phonemic level.

Pitt and Szostak (2012) demonstrated that the effect of phoneme manipulation is reduced when participants’ attention is explicitly directed to the ambiguous phoneme, with participants less likely to categorise an ambiguous item as a “word” under such conditions than otherwise. We applied this paradigm at the sentence level to investigate whether disfluencies induce heightened attentional focus at a phonemic level. Specifically, we compared the impact of a phoneme manipulation on lexicality judgements with; (i) attentional focus, and; (ii) disfluency presence.

The first experiment investigated the effects of focus without disfluency. Participants (n=40) listened to sentences consisting of an uninformative context followed by a sentence-final target word. They were then asked to judge the lexicality of the final word they heard. The multisyllabic target words (‘Impressive’ & ‘Condition’) included a word medial phoneme that was altered along a 5 stage continuum from /s/ to //, shifting them from word to non-word. Participants saw either only ‘focused’ or only ‘unfocused’ written instructions before the listening task. In the focused condition participants were alerted to the position and sound of the phoneme that could be changed. These details were not present in the unfocused instructions. The second experiment (participants, n=40) introduced disfluency into the paradigm. The method and materials were the same but half of the trials included a filled pause (‘uh’) immediately preceding the sentence final target.

In the first experiment, we found a focus effect: participants were less likely to judge the target word to be an actual word in the focused than the unfocused condition. This suggests that the effects observed by Pitt and Szostak (2012) can be generalised to words appearing in a sentential context. In the second experiment, we found that the presence of a disfluency caused participants to be more likely to judge that the target word was an actual word. Although the analyses of our second study are preliminary, they suggest that rather than heightening attention to phonemic information, the presence of disfluency may in fact be detrimental at this level.

References:
ELABORATION OF TARGETS AND NON-TARGETS IN ONLINE SENTENCE COMPREHENSION

Philip Hofmeister (University of Essex) & Shravan Vasishth (University of Potsdam)
phofme@essex.ac.uk

Across a range of tasks, the extent of semantic processing often predicts memory efficiency and accuracy, e.g. random trivia facts are remembered better after additional causally-related information, word recall & recognition benefits from meaning-related processing vs. superficial processing, and reading times within sentences are faster at retrieval sites given syntactically & semantically complex vs. simple targets [1-3]. Such effects are broadly interpretable on the view that additional processing creates distinctive memory representations, "making these memories less susceptible to interference and/or providing more features that can be cued" [2].

One possibility that this view raises is that a memory target will contrast more with other stimuli, and hence be remembered better, if non-target stimuli receive more attention or processing. Such an idea is plausible if providing more detail about any discourse referent lowers the chances for confusion with other memory retrieval candidates. The experiment described here tests this idea. In our self-paced, moving window reading study, participants (n = 52) read sentences varying with respect to the syntactic & semantic complexity of a target NP [= NP2] and a preceding non-target NP [= NP1], via the presence/absence of two prenominal modifiers:

(i) The (brutal military) policeman arrested the (peaceful Buddhist) monk who a councillor in the capital city saved from being imprisoned and tortured.

Two critical regions were analyzed: the head noun of the relative clause subject [= NP3] + the subsequent word, and the verb + subsequent word (separate analyses for each word). Previous results led us to expect that complex target NPs would lead to faster processing times at the retrieval region than simple NPs. The key question here was how the complexity of NP1 affects downstream processing, particularly at the retrieval site.

We derived residual log reading times by regressing log reading times against logged list position and word length with participant intercept adjustments. In the final mixed effects models, we included fixed effects for each complexity manipulation, their interaction, accuracy and interactions with the complexity variables, and reading times at the two preceding words, random effects for participants and items, and the maximal random effect structure. All analyses were carried out via Bayesian hierarchical models via rstan. At each of the two words in the retrieval region, we observed a main effect of target complexity, but no effect of non-target complexity, nor an interaction. At the encoding region for NP3, however, we found an interaction between the complexity variables: reading times were faster when at least one NP was complex, but markedly slower when all NPs were simple.

In sum, encoding additional syntactic and semantic features causes faster reading reading at retrieval sites, but only for downstream targets, whereas the effects tied to non-targets are restricted to encoding interference. Such findings argue against the idea that elaboration effects in sentence processing result from a general preference for descriptiveness, and the idea that greater all-around unique features or elaboration modulates the effect of similarity-based interference.

Participants exploit dependencies hidden in studies employing Latin Square design
Jakub Szewczyk (Jagiellonian University)
jakub.szewczyk@gmail.com

Latin Square designs are commonly used to counterbalance psycholinguistic experiments. They enable the researcher to recycle the same stimulus across different conditions, and treat the corresponding variable as both within-subject and within-item, despite the fact that each subject sees each item only once. For example, a subset of stimuli used in such an experiment could include following items:

Participant A
1. He spread the warm bread with butter.
2. The bill was due at the end of the month.

Participant B
1. He spread the warm bread with month.
2. The bill was due at the end of the butter.

In many experiments (as in above examples) target words are predictable given the preceding sentences. This exposes a dependency which can be exploited by participants. For example, after reading sentence 1 participant B could remember that it ended with the incongruent word month, instead of the expected word butter. When she then processes sentence 2, she can realize that it should be continued with month. However, since month has already been used up, she will know that this sentence is in the incongruent condition, even before encountering the target word.

Below we present an ERP study demonstrating that participants indeed exploit such dependencies. Participants were presented with short stories. All the stories were displayed on the screen all-at-once, except for the final sentence, which was presented word-by-word. In half of the stories, a target word (the direct object noun of the main clause of the story-final sentence) was semantically congruent with the preceding context (congruent condition [C]), while in the other half it was not (incongruent condition [IC]). The incongruity was introduced by reassigning target nouns from a given story to another story. In addition to the congruity manipulation, in half of stories, just before the story-final sentence, explicit information was introduced, telling the participants that in the following sentence one of two specific target words will appear (Expectancy condition [E]). The two words displayed were the congruent and the incongruent target word (participants did not know which of the two words will appear). In the other half of items, no information concerning the target word was given (No Expectancy condition [NE]). Expectancy and Congruity manipulations were fully crossed, resulting in four experimental conditions: NE-C and NE-IC, E-C, and E-IC.

The analysis of ERPs elicited by the words of the story-final sentence preceding the target word presentation (hereafter called intervening words) revealed that they were processed differently in the E-C and E-IC conditions, even though up to the target word items were identical in both of the conditions. This indicated that participants must have made strategic between-item predictions concerning the items, along the lines we described above. Three kinds of evidence support this conclusion: 1) For each participant, based on his/her individual history of items, we split items into those where the participant could have guessed item congruity based on prior items, and those where he/she could not. It turned out that the difference between the E-C and E-IC conditions at the intervening items was carried exclusively by the predictable items. Critically, the standard N400 effect for target nouns in the NE-IC condition (relative to NE-C condition) was smaller for predictable items, indicating that the effect pertains also to no-expectancy items. 2) A similar pattern of effects was obtained after splitting the participants into those with high and low episodic memory: the effect was carried primarily by participants with good recall of stimuli. 3) Finally, the effect was carried by stories that were semantically highly constraining at the target word position. These results indicate that using Latin Square designs with predictable stimuli can invite unwanted strategies for participants with good episodic memory.
ENGLISH RESUMPTIVE PRONOUNS ARE PREDICTED BY BARE-BONES SYNTAX

Adam Morgan (University of California San Diego)
a1morgan@ucsd.edu

Wh- dependencies are formed when a constituent appears in a clause-peripheral position as opposed to its canonical/base position. Across languages, there are two ways the base position can be realized: either phonologically null (gap strategy, 1a), or with a pronoun (Resumptive Pronoun strategy, 1b) [6]. In English, RPs were long held to be a rescue strategy for syntactically complex or illicit structures [2,10], but studies have since shown that RPs are never judged acceptable [1,2,5,9]. In production tasks, however, RPs are reliably elicited in certain syntactic contexts [3,7]. This pattern of data is puzzling: Why should English speakers have intuitions about a grammatical strategy from other languages, but which is never acceptable in their own? A recent study demonstrated that RP production rate is inversely proportional to the acceptability of the corresponding gap structure, leading to the claim that overall gap acceptability predicts RP rates [7]. The present studies test this hypothesis by varying stimulus acceptability with clausal and lexical manipulations.

Experiment 1. 44 participants rated the acceptability of 36 Wh- dependencies on a 1-9 Likert scale. 3 clause types (doubly embedded, Wh- complement, adjunct) known to span a range in acceptability [1,2,4,7,8,10] were fully crossed with 2 levels of memory load (high, low) and 2 Wh- dependency strategies (gap, RP). Memory load was manipulated by altering referential and syntactic properties of DPs to increase (2a)/decrease (2b) mutual distinctiveness [5], thereby decreasing (2a)/increasing (2b) memory load. The expected effect of clause type emerged (F(1,1255)=23.18, p<.001); crucially, increased memory load led to significantly lower ratings (F(1,1255)=20.21, p<.001). Therefore, if gap acceptability does predict RP rates, increased memory load should lead to higher rates of RPs in Exp.2.

Experiment 2. In an elicited production task, 31 participants typed in a textbox to complete a Wh- dependency (3b) based on information from a preceding sentence (3a). Responses were coded as ‘gap’ (3c), ‘RP’ (3d), or ‘other’ (passives, errors, etc.). The data show the expected effect of clause type (X²(2)=282.974, p<.001), but no effect of memory load on the proportion of RPs produced (X²(1)=.313, p=.576). This null effect is not likely to reflect insufficient power given the significant effects of memory load reflected in decreased gap production (z=-4.221, p<.001), a corresponding increase in ‘other’ responses (z=5.367, p<.001), and lowered ratings in Exp. 1.

Discussion. The insensitivity of RP production rate to memory load is problematic for views of resumption as a strategy for alleviating processing burdens [1,2,4,6,8,9,10]. Given the significant effect of memory load on acceptability (Exp. 1), the lack of effect in Exp. 2 also contradicts the hypothesis that it is the overall acceptability of the gap structure that predicts RPs [7]. A more accurate characterization may be that each clause type comes with a particular rate of resumption. That rate would correlate (possibly causally) with the acceptability of a bare schematic representation of the clause with a gap in it.

These data have interesting implications for sentence planning: The choice to produce a RP must occur before consideration of episodic details like lexicalizations, but after consideration of bare-bones syntactic information (given the predictive capacity of clause type for RP production). This implicates a stage in sentence production during which abstract structural schemas are planned, independent of lexical content.

1 a. the candidate that we campaigned for
   b. the candidate that we campaigned for her

2 a. That’s the orderly who somebody learned when Dr. House harassed __/her.
   b. That’s the orderly who the doctor learned when the nurse harassed __/her.

3 a. The doctor learned when the nurse harassed the orderly.
   b. That’s the orderly who the doctor filed a complaint when ____________.
   c/d. the nurse harassed __/him, her, them}

Relative clause extraposition in German was investigated in a production experiment combined with a reading span task and in a magnitude estimation experiment. Corpus studies show that rate of extraposition declines with increasing extraposition distance. For example, extraposition has to cross two words in (1) (*Geschenke geben*).

(1) Max wird dem Mann, (der das Haus baut), Geschenke geben, (der das Haus baut)
Max will the man who the house builds presents give who the house builds
‘Max will give the man who is building the house presents.’

Distance can be defined as number of words (Hawkins, 2004; Gildea and Temperley, 2010) or as number of new discourse referents (Gibson, 2000). In order to decide between these definitions of distance, 39 students participated in an oral production experiment which was a variant of the production-from-memory task. Participants first read a main clause as in (2). After a visual prompt like Max said that, the initial main clause had to be repeated orally from memory in the form of an embedded clause. While the initial main clause fixed the lexical content of the to-be-produced embedded clause, participants were completely free with regard to the position of the relative clause.

(2) a. Gratulieren wollte Max dem Lehrer, der gestern zu Besuch war.
   congratulate wanted P. the teacher who yesterday at visit was
   ‘Max wanted to congratulate the teacher who came for a visit yesterday.’
   b. (Einige) Gedichte vorlesen wollte Max dem Lehrer, der gestern zu Besuch war.
   Some poems read wanted P. the teacher who yesterday at visit was
   ‘Max wanted to read (some) poems to the teacher who came for a visit yesterday.’

The experiment varied the amount of material that had to be crossed by extraposition in addition to the verb: nothing (2-a), a bare NP object (2-b), or an NP object containing a determiner (2-b). The latter two conditions differ in number of words but are identical in number of new discourse referents. The results (see Table 1, % extraposed) show a drop in extraposition rate in the presence of an object but no significant difference between one- and two-word objects. This suggests that distance is defined as number of new discourse referents, not number of words. A magnitude estimation experiment with the same material revealed, however, that number of words had a small modulating effect on acceptability (see Table 1, z-score).

While the relationship between working memory and language comprehension has been the subject of numerous studies, few research exists concerning the involvement of working memory in language production (e.g., Hartsuiker and Barkhuysen, 2006; Kellogg, 2004). In order to investigate how working memory capacity affects word order choice, the reading span of each participant in the production experiment was determined on a scale ranging from 0 to 75 following the procedure recommended by Unsworth et al. (2005). The mean reading span was 48 (range: 27–70). The results show that the probability of extraposition correlates negatively with reading span (see Table 2; r = -.32, p < .05): participants with low reading span extraposed more often than participants with high reading span. Thus, working memory capacity as measured by the reading span task affects word order choice during language production, suggesting that language production is subject to the same kind of working memory constraints found for (certain aspects of) language comprehension. Furthermore, the finding of a negative correlation suggests that speakers extrapose primarily for their own sake and not for the sake of the hearer (see Wasow, 1997).
Theoretical linguists have debated the relationship between language and music (Jackendoff’06, Patel’08, Katz/Pesetsky’11), but few psycholinguistic studies have tested whether structural aspects of music prime structural aspects of language. If music and language have shared representations, we expect priming from music to language. Initial evidence comes from Van de Caevey/Hartsuiker’11, who used sequences of sine tones differing in pitch, and found priming on Dutch relative clause (RC) attachment, and Menon/Kaiser’13, who used piano notes and found priming of English RC attachment.

Our study has two main aims: First, we investigate whether musical notes played on a piano (easily recognized as part of the domain of music) trigger priming in the linguistic domain. Second, we test whether potential priming effects interact with the singular/plural distinction on nouns (Fernández/Sainz’04). Researchers disagree on whether Spanish RCs have default high or low attachment (Carreiras etal’01, Cuertos/Mitchell’88); Fernández /Sainz’04 suggest that whether nouns are singular or plural influences attachment patterns. We tested this in production and looked at whether it interacts with priming, thereby testing the ‘primability’ of a language whose attachment bias differs from English.

EXPERIMENT (n=20, 30 targets, 120 fillers): We primed participants with melodies created using the ‘Circle of Fifths’ where tones from musical keys combine to form chords/chord progressions; after hearing a prime, participants (Spanish speakers in Spain) wrote a continuation for an RC fragment. MUSICAL PRIMES were 8-note piano melodies. We manipulated (a) attachment height (high/low) and (b) pause (pause/no-pause). In high-attachment (HA) primes, the last two notes connect to the initial harmonic domain. In low-attachment (LA) primes, the last two notes connect to the second harmonic domain (figure). We also had a baseline melody condition with no structure. Durations and intervals of notes were constant (but pause conditions had a pause before two final notes).

LINGUISTIC TARGETS: People completed RC fragments with two nouns differing in number, e.g. Bob saludó al líder de los activistas que ‘Bob greeted the leader of the activists who’ (N1 Singular+N2 Plural) or Tina conoció a los jardineros del millonario que ‘Tina knew the gardeners of the millionaire who’ (N1 Plural+N2 Singular). We analyzed whether continuations modify the complex NP “the leader of the activists” (HA) or the lower NP “the activists” (LA).

RESULTS: There is a significant effect of target type, Plural+Singular vs. Singular+Plural (p<.05), on participants’ attachment choices: N1 Plural+N2 Singular targets show an overall LA bias in the baseline condition (63% LA), which fits with Fernández/Sainz’04 (who didn’t test priming). Furthermore, these targets show no significant priming (no effect of attachment height or pause, no interaction), perhaps due to a ceiling effect: All conditions have more LA than HA completions. In contrast, N1 Singular+N2 Plural targets show no clear attachment biases in the baseline condition (50% LA), similar to Fernández/Sainz’04. Crucially, in contrast to the Plural+Singular targets, the Singular+Plural targets do show priming from music: The rate of HA continuations is significantly greater after high-attachment primes (p<.05). There is no effect of pause and no interaction.

Our baseline results support Fernández/Sainz’04’s finding that the singular/plural distinction influences RC attachment (cf. Fernández/Sainz’04 on processing load of plurals possessors). Our priming results suggest that explicitly musical stimuli can activate abstract representations that overlap with the syntactic representations of language, but that cross-domain priming effects are modulated by the singular/plural distinction and associated RC attachment biases. Furthermore, the finding that a pause after the second noun does not boost HA rates (cf. Fodor’98) suggests rhythmic cues are not needed for priming, indicating that what matters is the structure of the harmonic domains. IN SUM, our results point to a domain-general level of abstraction in the representation of structural information.
PARAFOVEAL PREVIEW BENEFIT DURING SECOND LANGUAGE READING
Chi Yui Leung (Nagoya University) & Masatoshi Sugiura (Nagoya University)
sieileung@nagoya-u.jp

Introduction. Parafoveal processing during silent reading has been well investigated among first language (L1) readers, but not among second language (L2) ones. Since parafoveal processing is a function of processing difficulty [1] and L2 reading generally incurs a higher processing cost [2], we anticipated that the parafoveal preview benefits for L2 readers would be different from those for L1 readers. Empirical evidence, however, is lacking at present. This study aimed to examine if and how parafoveal preview benefits vary between L2 and L1 readers by using transposed-letter (TL) non-words in parafoveal previews with a gaze-contingent boundary paradigm [3].

Method. Forty L2 readers (Japanese learners of English; mean TOEFL PBT score: 533) and 14 L1 readers were eye-tracked (EyeLink 1000) while reading 126 sentences (including 42 fillers). The L2 readers were further divided into advanced (AdvL2) and intermediate (IntL2) groups using a median split for their TOEFL scores so as to examine the effect of L2 proficiency. Eighty-four five-letter words (e.g. train) embedded in sentences (e.g., We took a fast train to Tokyo yesterday) were used as target words. Parafoveal previews of the target words were divided in four conditions: (1) Identical (the target word itself; e.g., train), (2) TL-Internal (transposing the 3rd and 4th letters of the target word; e.g., trian), (3) TL-Initial (transposing the 1st and 2nd letters of the target word; e.g., rtain), and (4) Unrelated non-words (all letters were consonants and none were part of the target word; e.g., qsvml). An invisible boundary was set at the end of the pretarget word (e.g., fast). Apart from the eye-tracking experiment, a lexical decision task with the TL stimuli as non-words was performed. This showed that the RTs for judging the TL-Initial items as non-words were shorter than those for judging the TL-Internal items for both the L2 and L1 readers.

Results. Eye-movement data were analyzed using linear mixed-effects models with participants and items as random effects. Group (L1 = baseline), Preview (Unrelated = baseline) and the Group × Preview interaction were treated as fixed effects. The model for log-transformed gaze durations (GDs) on the target words revealed significant interactions, and each participant group was further analyzed separately. The results showed that GDs for the Identical, TL-Internal and TL-Initial conditions were significantly shorter than those for the Unrelated for both the L1 and AdvL2 groups. However, only the Identical and TL-Initial but not the TL-Internal conditions yielded shorter GDs than the Unrelated condition did for the IntL2 group (see figure).

Additional modeling including incoming saccade amplitude (ISA) and GDs on the pretarget words (preGDs) as covariates revealed that GDs for the Identical, TL-Internal, and TL-Initial conditions were shortened by either decreasing ISA or increasing preGDs (relative to the Unrelated condition) for the L1 group. In contrast, these kinds of effects from preview space/time [4] were either absent for the AdvL2 group or worked in the opposite direction in the TL-Internal condition for the IntL2 group (i.e., GDs were inflated by longer preGDs).

Conclusion. These results indicate that L2 reading is also facilitated by parafoveal preview benefits. Interestingly, our data suggest that there may be qualitative differences between the parafoveal processing of L1 and L2 readers, as well as among L2 readers of different L2 proficiencies. These are discussed in terms of "preview cost" [4] and lexical quality [5].

NO MORPHOLOGY IN THE PARAFOVEA: EVIDENCE FROM AN INVISIBLE-BOUNDARY EYE-TRACKING EXPERIMENT

Federica Degno (University of Southampton) & Davide Crepaldi (University of Milano Bicocca)
davide.crepaldi1@unimib.it

Masked priming data suggest that complex words are decomposed into morphemes not only when the entire word shares a semantically transparent morphological relationship with its stem (e.g. *prigioniero-prigione*, prisoner-prison), but also when the word can be simply segmented orthographically into two existing morphemes (e.g. *ostaggio-oste*, hostage-host; Davis and Rastle, 2010). This implies that early morphological analysis is triggered by morpho-orthography rather than morpho-semantics. However, there is no clear demonstration that orthography-driven morphology is a factor during natural sentence reading. In this study, we address this issue in a gaze-contingent, eye-tracking study where complex words were manipulated in the parafovea.

96 university students were asked to read silently and understand 36 sentences that contained a target word. Target words were either genuine derived words, i.e., they shared a semantically transparent morphological relationship with their stem (e.g. *prigioniero-prigione*, prisoner-prison); or opaque words, i.e., they had no semantic relationship with their (pseudo-)stem, but also included a pseudo-suffix, and thus were entirely parsable morphologically (e.g. *ostaggio-oste*, hostage-host); or orthographic controls, i.e., simple words made up of an existing stem and a non-morphological ending (vitittima-vitto, victim-board; an analogous example in English would be dialog-dial). As in the classic gaze-contingent paradigm, an invisible boundary was placed at the beginning of the blank space before the target words, and target words appeared either in their correct form (valid preview condition), or with their suffix/pseudo-suffix/non-morphological ending substituted by random letters (e.g. *prigioniero-prigionlapc*) before participants crossed this boundary with their eyes. Critically, target words in the three conditions were arranged in triplets that could fit the same sentence, so that any effect related to the sentence context was perfectly balanced across conditions. We also took care that cloze probability in those sentences was matched between transparent, opaque and control words. In addition, we asked participants to fill a questionnaire at the end of the experiment in order to control whether they were aware of the display change.

Parafoveal-on-foveal effects were generally weak or absent. On the contrary, single fixations and gaze durations on the target word were significantly shorter after valid previews. Critically for the purpose of the study, none of these effects interacted with the morphological structure of the target word, that is, the preview benefit effect was equivalent in transparent, opaque and non-morphological target words (see Table 1). This suggests that words in the parafovea are not processed morphologically, which shows some clash between models of sentence reading (e.g., Risse & Kliegl, 2014) and models of the visual identification of complex words (e.g., Rastle, Davis & New 2004).

Table 1. Single fixation durations, means and standard errors of the mean in ms.

<table>
<thead>
<tr>
<th></th>
<th>Transparent</th>
<th>Opaque</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invalid preview</td>
<td>254 (SEM=8.9)</td>
<td>251 (8.8)</td>
<td>247 (8.2)</td>
</tr>
<tr>
<td>Valid preview</td>
<td>245 (9.0)</td>
<td>243 (8.2)</td>
<td>236 (8.4)</td>
</tr>
</tbody>
</table>

References
EFFECTS OF BEAT GESTURE AND PITCH ACCENT ON THE PROCESSING OF WORDS IN CONTEXT

Diana Dimitrova (Radboud University Nijmegen & Donders Centre for Cognitive Neuroimaging), Mingyuan Chu (MPI for Psycholinguistics), Lin Wang (Chinese Academy for Sciences), Asli Özyürek (Radboud University Nijmegen & MPI for Psycholinguistics) & Peter Hagoort (MPI for Psycholinguistics & Donders Centre for Cognitive Neuroimaging)
Mingyuan.Chu@mpi.nl

Speech is organized according to information structure: important information (focus) is highlighted and distinguished from less important information (non-focus). Previous studies suggest that listeners process accented information more elaborately. Information can be emphasized non-verbally too, for instance by beat gestures, which are rhythmic hand movements without semantic meaning (McNeill, 1992). ERP studies have shown that beat gestures facilitate phonological, syntactic and semantic processing (Biau, & Soto-Faraco, 2013; Holle et al., 2012; Wang & Chu, 2013). The questions arise whether beat gestures and pitch accents show a similar neural signature due to their similar functionality, and whether listeners expect beat gestures to be aligned with the focus of the message.

In the present ERP study we investigated how beat gestures modulate the processing of accented-focused vs. unaccented-non focused words in context. 29 Dutch native speakers watched movies consisting of short dialogues (see examples below). In the accent condition (1), the target was accented and in focus, and it was accompanied by (i) a beat gesture, (ii) a control hand movement (self-touching movement) or (iii) no gesture (speaker stood silently). In the no accent condition (2), the target was unaccented and not in focus and was combined with the same gesture conditions. Gestures started 520 ms prior to target word and reached their maximal extension at the onset of the target. All gesture parameters (hand shape, naturalness, duration, alignment) were tested in behavioural tests.

ERPs were time-locked to the gesture onset and analyzed. We performed a cluster-based random permutation tests to test for main effects and interactions of the factors. We found a main effect of pitch accent at 300 ms post target onset: accented words elicited a positivity relative to unaccented words. Gesture triggered a main effect too. Words accompanied by a beat gesture or by a control movement elicited a frontal positivity and a centro-parietal negativity at gesture onset (0-300 ms), relative to words without a gesture. The gesture effect continued as a positivity (300-800 ms). After 800 ms, control movements elicited a negativity relative to words without a gesture. We found interactions of each gesture condition and pitch accent. Accented words with a beat gesture elicited a positivity relative to accented words without a gesture (300-800 ms post target onset). No positivity was found in the same time window for unaccented words with vs. without a beat gesture. Accented words with a control movement triggered a positivity and a negativity relative to accented words without a gesture (300-800 ms post target onset). Unaccented words with a control movement elicited only a positivity relative to unaccented words without a gesture.

Our findings are in line with the ERP literature (Wang & Chu (2013). Our data indicate that listeners integrate beat gestures with speech and do not perceive beats on non-focused information as a mismatch. The negativity for control hand movements suggests that they may increase processing difficulty relative to words without gestures. We conclude that beat gestures play a unique role in the processing of accented words in context, presumably because more attentional resources are allocated at focus information in context.

Examples of experimental stimuli

1. Q: Did the student buy the books or the magazines via Amazon?
   A: He bought the BOOKS via Amazon.

2. Q: Did the student buy the books via Amazon or via Marktplaats?
   A: He bought the books via AMAZON.

Note: Target words are underlined, accented focused words are in capitals.
ONLINE PROCESSING OF MORAL TRANSGRESSIONS: ERP EVIDENCE FOR SPONTANEOUS EVALUATION

Hartmut Leuthold (University of Tuebingen), Angelika Kunkel (University of Tuebingen), Ian G. Mackenzie (University of Tuebingen) & Ruth Filik (University of Nottingham)
hartmut.leuthold@uni.tuebingen.de

Recent accounts of moral judgment have rejected the long-held assumption that moral judgments reflect the output of a rational reasoning system. Specifically, experimental studies using fictional moral dilemmas have indicated that both automatic emotional processes and controlled cognitive processes contribute to moral judgments. Whereas these studies employ vignettes that describe the moral scenario and subsequently ask participants to make an explicit moral decision, the actual online comprehension processes during reading of the materials and the associated implicit moral evaluation are not considered. In addition, the materials often refer to moral dilemma situations that are rarely encountered in the social environment (e.g., killing a person to rescue others). As a result, not much is known about how people process socio-normative violations that are more common to their everyday life, nor the type and time-course of the involved online processes.

The aims of the current research are therefore to investigate on-line evaluative processes that are spontaneously triggered by vignettes (in German) that describe more probable situations. To this end a text comprehension paradigm was employed that did not demand explicit evaluative judgments. We created prototypical scenarios (see below) in which an initial sentence was used to establish the social context. The test sentence contained a critical word that described a behavior that was either morally acceptable or unacceptable relative to the context. To assess evaluative processes, we recorded participants’ (N = 28) electrical brain activity (ERPs) while they were reading the vignettes. In order to check for the specificity of ERP effects associated with moral transgressions, we included additional materials that differed with respect to a morality-unrelated text dimension, containing information that was either consistent or inconsistent with the participants’ general world knowledge. The materials for the resulting four experimental conditions were as follows (only the target sentence is presented in German together with its word-by-word translation, disregarding word order in English; the critical word is in italics):

Morally-acceptable context: Tina’s grandfather suffers from cancer and will die soon. For his 85th Birthday he has planned a big party and wishes nothing more than that all his children and grandchildren attend.

Morally-unacceptable context: Tina’s boss has been making explicit innuendos for some time. Now he has invited her out for dinner to an expensive restaurant. She knows that he has been married for 20 years and is the father of three children.

Target sentence: Sie hat die Einladung angenommen. [She has the invitation accepted.]

Knowledge-consistent: During a France exchange Mrs. Lehmann eats a famous French speciality.

Knowledge-inconsistent: Mrs. Lehmann goes to a Schwabian Restaurant and orders a local speciality.

Target sentence: Sie erhält als Gericht einen Teller voller Schnecken und Weißbrot. [She receives as dish a plate full of snails and white bread.]

As expected, ERPs showed a larger centroparietal negativity for world knowledge-inconsistent words about 300 ms after critical word onset (N400), indicating an influence on context-dependent meaning construction. By contrast, at around the same time after word onset, morally unacceptable compared to acceptable words elicited a larger ERP positivity. We take this early ERP positivity to reflect an implicit evaluative (good-bad) categorization process that is rapidly engaged during the online processing of moral transgressions. The absence of an N400 effect might be attributed to the less constraining moral contexts, reducing the demands on meaning construction in the case of inconsistencies.
EXAMINING THE NEURAL CORRELATES OF IRONY AND SARCASM: AN FMRI STUDY
Ruth Filik, Alexandra Turcan, Christina Ralph-Nearman & Alain Pitiot (University of Nottingham)
ruth.filik@nottingham.ac.uk

Despite the fact that irony and sarcasm are very common in everyday language use, little is known about the brain areas that are involved in their processing. The conventional view of verbal irony is that it is a figure of speech that communicates the opposite of what is said, for example, by saying “What lovely weather” in the middle of a storm. Sarcasm is a form of irony that, as well as communicating the opposite of what is said, is generally intended to be critical or hurtful to the recipient (e.g., McDonald & Pearce, 1996), for example, uttering the sarcastic comment, “That was clever”, when the recipient has done something stupid. Thus, although both irony and sarcasm involve computation of what the speaker actually intends to say, it seems likely that the comprehension of sarcastic irony also involves more of a socio-emotional component. Therefore, while both sarcastic and non-sarcastic irony might activate areas involved in mentalising (e.g., Spotorno et al. 2012), sarcastic irony might also activate social or emotion-related brain areas (e.g., Uchiyama et al., 2012).

In the current study, we used functional magnetic resonance imaging (fMRI) to investigate the neural correlates of sarcastic and non-sarcastic irony, compared to literal language. Seventeen participants read scenarios in which the final comment was either intended literally, as non-sarcastic irony (i.e., not intended literally, and not intended to criticise the recipient), or as sarcastic irony (i.e., not intended literally, and intended to criticise the recipient):

Literal comment: Bernice and Caitlin were both applying for a Psychology course at a university in the USA. They went to print out their applications together. Caitlin chose to print hers on letter-headed paper. Bernice said to Caitlin: “Very formal!”

Non-sarcastic irony: Bernice and Caitlin were both applying for a Psychology course at a university in the USA. They went to print out their applications together. The printer only had pink paper available. Bernice said to Caitlin: “Very formal!”

Sarcastic irony: Bernice and Caitlin were both applying for a Psychology course at a university in the USA. They went to print out their applications together. Caitlin chose to print hers on pink paper. Bernice said to Caitlin: “Very formal!”

The fMRI and anatomical scans were acquired on a Philips 3T MR scanner. A mixed-effect, event-related statistical analysis of the signal change associated to the final comment was performed using SPM8’s standard two-level general linear model approach. In terms of key findings, the Non-sarcastic irony > Literal contrast showed greater activity in Brodmann areas 6 (left), 13 (left and right insula), 19 (right), 22 (right) and 39 (right). These areas are not only associated with language processing but also with Theory of Mind (ToM), the involvement of which in the understanding of irony has recently been demonstrated (Spotorno et al., 2012). Interestingly, BA 39 is linked to the understanding of the relationship between characters in reading tasks, and the insula has been implicated in humour appreciation (Moran et al., 2003), which would support the notion that one function of irony is to be humorous (e.g., Bowes & Katz, 2011). In relation to the processing of sarcastic irony, in line with our predictions, the Sarcastic irony > Literal contrast showed activation in BA24 (right) and BA31 (left), which are associated with emotional processing, in addition to BA21 (right) and BA40 (left), which are associated with text processing.

In conclusion, the results suggest that, as predicted, the comprehension of irony involves the computation of the speaker’s attitude and intended meaning; reflected by the activation of the ToM network. This is in line with cognitive theories of irony (e.g. Echoic Mention Theory), which propose that an essential part of irony comprehension is a correct identification of the speaker’s attitude and intention. Furthermore, it appears that sarcastic irony (but not non-sarcastic irony) requires an additional emotional processing of the comment, reflected in the activation of parts of the brain’s emotional network.
MODELING NEURAL CORRELATES OF SYNTACTIC STRUCTURE BUILDING

John Hale, Frederick Callaway, Elana Feldman, Jaclyn Jeffrey-Wilensky, David Lutz, Adam Mahar (Cornell University) & Jonathan Brennan (University of Michigan)

jthale@cornell.edu

Using fMRI, we find that the anterior temporal lobe is selectively involved in a computation whose resource consumption profile is isomorphic to that of an incremental phrase-structure parser. This finding extends Brennan et al. (2012) using American English and syntactic analyses grounded in current linguistic theory.

We scanned 11 right handed college-age participants who reported no neurological abnormality. Over headphones, these participants listened to the first chapter of Lewis Carroll’s Alice in Wonderland, as read by a speaker of American English (12 m 23 s).

We used the text of the story to derive time series predictors, and fit these predictors to preprocessed fMRI data in order to identify regions with correlated activation. One of these predictors simply marks the offset of each spoken word. This predictor localized regions in temporal lobe whose BOLD time series correlates with the rhythm of the spoken narrative. Prior work suggests that the anterior temporal lobe is involved in "basic syntactic processes" (Friederici and Gierhan 2013). We then evaluated the ability of text-based syntactic predictors to account for mean BOLD signal in a 10mm spherical region around these participant-specific peaks.

These text-based predictors were all based on the number of syntactic nodes added to the parse tree word-by-word. Following Van Wagenen et al. (2012), these predictors differed in terms of whether the parsing algorithm was (T)op-(D)own or (B)ottom-(U)p, and in terms of whether the underlying grammar was based on the (P)enn Treebank (Marcus et al., 1993), or (M)inimalist Grammars (Stabler 2013). We convolved these node counts with a canonical hemodynamic response function and then orthogonalized them with respect to the word-offset predictor used for localization. We used the resulting time series’ to predict BOLD signal in the anterior temporal regions of interest described above.

Using linear mixed effects models, each node count predictor was significant at the $p < 0.05$ level for the left anterior region of interest (BUP $\beta$=0.0246 SE=0.0124; TDP $\beta$=0.0280 SE=0.0123; BUM $\beta$=0.0458 SE=0.1216; TDM $\beta$=0.0559 SE=0.1212). They were not significant in the homologous right hemisphere region. Comparing grammar types and prediction directions, the best fit was achieved using structures generated by Minimalist Grammars with a Top-Down parsing regime (BIC BUP = 9233; TDP = 9231; BUM = 9223; TDM = 9216).

This result represents a kind of match between the sequence of abstract operations taken by an incremental parser and observed BOLD signals in a region that has been assigned to basic syntactic processes.


RECASTING THE ELAN AS AN ATTENTIONAL EFFICIENCY-DEPENDENT ENHANCEMENT OF THE N100
Christopher Barkley, Robert Kluender, & Marta Kutas (University of California San Diego) cbarkley@ling.ucsd.edu

Language researchers have long been interested in the nature and extent of the language system's interfaces with domain-general cognitive systems. The current study was designed to investigate the relationship between language and attention, which has been neglected due to an emphasis on the language-working memory interface [1]. Here we investigate this relationship using the eLAN, an early (100-300 msec) left anterior negativity originally elicited in [2] by comparing the bolded “word-category violation” (WCV) to its control, and subsequently interpreted as an index of first-pass structure-building operations [3].

The man admired {a sketch/*Don's} OF sketch the landscape.

Rather than a response indexing linguistic structure building operations [3] or low-level sensory form-based processing [4], we hypothesized that the eLAN is ontologically an attentional response. While we agree with the functional interpretation in [4], it lacks a mechanistic explanation as to why WCVs elicit eLAN.

A review of the attention literature suggests that an attentional account can provide such a mechanistic explanation. WCVs by hypothesis engage the executive attentional system when language users monitor for ungrammaticality, orienting them to the unexpected, task-relevant stimulus [6], resulting in selective attention to the WCV, enhanced sensory processing, and an increase in the amplitude of the N100 [7]. The efficiency of these attentional processes in other contexts is known to be highly variable across individuals, as demonstrated by variability in performance on the Attention Network Task (ANT) [8], which yields estimates of the efficiency of the brain's attentional (a) alerting, (b) orienting, and (c) executive networks [6]. We hypothesize that the efficiency with which the brain responds to linguistic input, here indexed by the eLAN, is dependent on the efficiency of these attention systems, and that the physical parameters of the eLAN should co-vary with the efficiency of the executive attention and/or orienting networks.

Methods: We recorded ERPs and behavioral responses while 36 subjects completed both the ANT and a sentence processing experiment containing WCVs and their controls (see above; 40 per condition), embedded in 240 filler sentences similar to those in [1].

Results: Behavioral ANT results were consistent with the literature in demonstrating (i) independence of an individual's attention networks (all $R^2 < .23$), and (ii) variability in the efficiency of attentional networks within and across participants. eLAN effects were also highly variable across individuals. We conducted median split analyses that divided subjects into high and low efficiency (alerting, orienting, and executive) groups (n=16) based on ANT performance. Significant eLAN effects were observed in two groups only: those with low-efficiency orienting systems ($p < .05$) and those with low-efficiency executive systems ($p < .05$), and importantly not in their high-efficiency counterparts (all $p > .5$). Crucially, eLAN effects were present or absent across attentional efficiency groups, whereas LAN, N400, and P600 effects were present in all groups, and showed only minimal differences in latency, amplitude, and scalp distribution across groups. Results of regression-based analyses will also be discussed.

Conclusion: These data suggest that eLAN effects may be mere attentional modulations of the N100. We hypothesize that while comprehenders with high-efficiency attentional systems possess adequate resources to accommodate WCVs, low-efficiency comprehenders must engage additional selective attentional resources in order to process WCVs, increasing N100 amplitude. This finding highlights the importance of investigating cognitive systems beyond working memory in sentence processing contexts.

From case markers to phrase structures: ERP evidence on predictive sentence processing in Korean
Yunju Nam (Konkuk Univ.), Hongoak Yun (Konkuk Univ.), Eunyong Chung (Konkuk Univ.) & Upyong Hong (Konkuk Univ.) supia05@konkuk.ac.kr

It is well known that processing difficulties caused by the limitation of cognitive capacities could be reduced either by expectation-driven processing strategies (expectation-based approach, EbA, Levy, 2008) or by strategies for cutting down integration/memory cost (memory-based approach, MbA, Gibson, 1998). This study presents ERP evidence which supports the former idea, by investigating the processing of center-embedded clauses in Korean, where double nominatives or double accusatives as well as scrambling are allowed:

Table 1. Sentence examples (e.g., (1) “Dongsu forgave Yunju whom Minho blamed”)

<table>
<thead>
<tr>
<th>Word Order</th>
<th>R1</th>
<th>R2</th>
<th>R3 (target)</th>
<th>R4</th>
<th>R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SOV</td>
<td>Dongsu-ka</td>
<td>Minho-ka</td>
<td>pinanha-n</td>
<td>Yunju-lul</td>
<td>yongse-hayss-ta</td>
</tr>
<tr>
<td></td>
<td>Dongsu-ACC</td>
<td>Minho-ACC</td>
<td>blame-EMB</td>
<td>Yunju-ACC</td>
<td>forgive-PAST</td>
</tr>
<tr>
<td>2 SOV</td>
<td>Dongsu-ka</td>
<td>Minho-lul</td>
<td>pinanha-n</td>
<td>Yunju-lul</td>
<td>yongse-hayss-ta</td>
</tr>
<tr>
<td></td>
<td>Dongsu-ACC</td>
<td>Minho-ACC</td>
<td>blame-EMB</td>
<td>Yunju-ACC</td>
<td>forgive-PAST</td>
</tr>
<tr>
<td>3 OSV</td>
<td>Dongsu-lul</td>
<td>Minho-lul</td>
<td>pinanha-n</td>
<td>Yunju-ka</td>
<td>yongse-hayss-ta</td>
</tr>
<tr>
<td></td>
<td>Dongsu-ACC</td>
<td>Minho-ACC</td>
<td>blame-EMB</td>
<td>Yunju-ACC</td>
<td>forgive-PAST</td>
</tr>
<tr>
<td>4 OSV</td>
<td>Dongsu-lul</td>
<td>Minho-ka</td>
<td>pinanha-n</td>
<td>Yunju-ka</td>
<td>yongse-hayss-ta</td>
</tr>
<tr>
<td></td>
<td>Dongsu-ACC</td>
<td>Minho-NOM</td>
<td>blame-EMB</td>
<td>Yunju-NOM</td>
<td>forgive-PAST</td>
</tr>
</tbody>
</table>

According to the EbA, double nominatives such as (1) or double accusatives such as (3) can signal the occurrence of a clause in the downstream of a sentence (about 60% from cloze task). That is, at R2 of (1) and (3), processors develop a specific expectation for an embedded clause, which would be met by a verb with an embedded clause marker -n at R3. In sentences like (2) and (4), in contrast, processors expect for a main verb to occur at R3 (about 95% from cloze task). Nevertheless, they encounter a verb with an embedded clause marker, which would lead to an extra processing load. Hence the EbA predicts that the processing of the verb forms at R3 would be easier in (1)/(3) than in (2)/(4).

According to the MbA, in contrast, processors would conveniently integrate the subject/object at R1 and the object/subject at R2 into the verb at R3 in (2)/(4), in order to minimize the memory load with having all constituents integrated. However, this is not the case for (1)/(3), since it is not possible for two consecutive nominative or accusative constituents to be fully integrated at R3. Thus, processors have to keep at least one not-yet-integrated constituent in their working memory. Consequently, the processing of the verb forms at R3 would be easier in (2)/(4) than in (1)/(3).

In an ERP experiment to determine which of the two scenarios is to be supported, we observed the P600 component in condition (1) compared to (2) at R2 and a long lasting late positivity effect in (2) compared to (1) at R3, although none of these ERP patterns was detectable in the comparison between (3) and (4). We interpret these results as evidence supporting the EbA: Processing difficulty increases if the incoming noun (R2) is marked with a less expected case (nominative after nominative, or accusative after accusative), and decreases if the incoming verb (R3) is in a highly expected form (a verb with a clausal suffix where an embedded clause is expected). In short, our results demonstrate that during sentence comprehension, Korean speakers actively use the information provided by case marker(s) to build expectations on how the syntactic structure of the sentence would unfold.

References
WORKING MEMORY INFLUENCES PROCESSING OF WHETHER-ISLANDS:  
THE N400 REFLECTS GAP PREDICTABILITY IN HIGH SPAN READERS  

Dan Michel (University of California, San Diego), Robert Kluender (University of California, San Diego) & Seana Coulson (University of California, San Diego).  
dcmichel@ucsd.edu

Reading time studies of syntactic islands [1] have suggested that the parser is immediately sensitive to island constraints on-line, and does not posit gaps inside islands [2,3]. But these studies have not examined what the parser actually does if it encounters a gap in an island. By testing the processing of island violations, we observe that many filler-gap-related ERP responses do not differ between islands and non-island controls. Both high and low span readers identify and fill gaps embedded in island and control sentences (evidenced by P600, LAN effects), but only high span readers show an additional effect inside of an island (N400, interpreted here as due to not expecting evidence for a gap inside an island).

(1a) Who had openly assumed [that the captain befriended the sailor before…
(1b) Who had openly inquired [whether the captain befriended the sailor before…
(1c) Who had the sailor assumed [whether the captain befriended _ openly before…
(1d) *Who had the sailor inquired [whether the captain befriended _ openly before…

Successfully interpreting (1c,d) requires associating the “gap” (underscore) with the displaced object who (“filler”). Filler-gap dependencies into interrogative subordinate clauses (“whether-islands,” 1d) are less acceptable than into declarative clauses (1c) [4,5]. In grammatical filler-gap dependencies, pre-gap verb positions (befriended) [7-9] typically elicit a P600, and LAN effects post-gap (before) [10,11], interpreted here as gap identification and filler-gap association, respectively. If gaps are more difficult to identify and/or associate with a filler in an island (1d), we expect to see this difficulty in the P600 and/or LAN response.

Methods: EEG was recorded from 29 electrodes as 32 monolingual English speakers (who completed reading span task) read 40 trials each of (1a-d) (+80 filler sentences) in RSVP format with 500ms SOA with a comprehension task. Noun phrases were matched with adverbs of equivalent mean frequency to control for sentence position (norming acceptability unaffected by presence/absence of adverbs). ANOVAs and paired comparisons were run on mean area amplitude in standard language ERP latency windows: 300-600ms (N400,LAN), 600-900ms (P600). Post-completion, acceptability ratings were obtained.

Results: Like prior findings ([5,6]), (1d) was rated least acceptable (p < 0.001). The results below do not reflect effects of syntactic satiation in these sentences [12].

i) befriended (pre-gap) P600 effect: (1c) > (1a) (p = 0.03), (1d) > (1b) (p = 0.005). Indicates gap identification in both grammatical (1c) and ungrammatical island (1d).

ii) before (post-gap) LAN effect: (1c) > (1a) (p < 0.001), (1d) > (1b) (p < 0.001). Indicates filler-gap association in both grammatical (1c) and ungrammatical island (1d).

Discussion: Our ERP results provide no evidence of reanalysis or a failed parse in (1d). They rather suggest that, after encountering an island boundary, high span readers lower their predictions for a gap inside an island, and are more surprised when one is encountered (N400 effect), indicating an influence of working memory capacity on the on-line processing of islands. Despite this, both high and low span readers rated the sentences the same offline, and showed the same gap-related brain responses in both islands and non-islands online (1c,d: P600 and LAN). Thus individual differences in the predictability of the presence of a gap do not prevent filler-gap identification/association in ‘ungrammatical’ whether-islands.

The Downstream Consequences of Form-To-Expectation Mismatch

Language comprehension arises by synthesizing stored knowledge about language with information available in a linguistic stimulus as it unfolds. Readers and listeners use linguistic, visual, and social contexts to generate predictions for many levels of the incoming input, from grammatical properties of a sentence [1], to semantic properties of a word [2], all the way down to low-level form-based perceptual properties of the physical input [3].

[4] demonstrated that when a sentential context conferred a strong expectation for a Noun (as in The child saved the … ex. 1a & 1b), subjects were slower to read the incoming noun when its physical form (i.e., its phonological/orthographic properties, quantified by a feature-based measure of word-form typicality [4]) was atypical (1b), as opposed to typical (1a), with respect to other words in the expected category. Furthermore, an MEG experiment [3] showed that at about 100ms post-stimulus onset— timing most directly associated with perceptual processing—a neural response was elicited when there was a mismatch between form and syntactic expectation. Moreover, the source of the effect was localized to the occipital lobe, consistent with the hypothesis that form-based syntactic expectations are influencing early perceptual processing of a recently encountered written word.

Here, in an ERP experiment, we investigate the degree to which a mismatch between a contextually-conf erred lexico-syntactic expectation and the low-level form-based properties of the incoming sensory information influence later-occurring responses indexing syntactic processing. 60 sentence frames (1) were constructed to contain nouns that had physical form-features that were either typical (1a) or atypical (1b) of other nouns (as per [4]), and were controlled for plausibility, frequency, # nearest neighbors, and length.

(1a) The child saved the marble that he … (Noun-like Noun)
(1b) The child saved the insect that he … (Verb-like Noun)

The specific word-form could not be predicted, but the category could. Hints of an early-occurring typicality effect were elicited, along with a statistically reliable P600 (see Figure 1). A stronger response occurred for atypical than for typical nouns. Many studies have demonstrated that the P600 can be influenced by orthographic irregularities [5]-[7]. Here, we show that well-formed words (appearing in perfectly grammatical sentences) that were perceptually atypical of other words in the same predicted category also elicited a P600 effect. The data complement results from [3], who demonstrated a robust effect of form-to-expectation mismatch early in processing. The present results illustrate the downstream effect of this mismatch at higher levels of signal analysis.

We propose that comprehenders use internally generated predictions at multiple levels to explain the source of the input [8]. A stream of hierarchically organized generative models propagates higher-level expectations to lower-level (progressively closer to perceptual cortex) models via feed-forward connections. At each level, as the incoming signal is intercepted, mismatch between input and expectation produces prediction error, which is then fed forward into the system. From such a perspective, form-to-expectation matching, and model updating based on error signals early in the processing stream, have rippling effects that permeate later processes. Form-to-expectation matching can play a central role in comprehenders’ ability to rapidly evaluate a linguistic signal.


Figure 1: Grand average ERP waveform at electrode site F3 to atypical verb-like (black line) versus typical noun-like (red line) nouns.
CLIQS: Cross-linguistic Investigations in Quantitative Syntax

Large-scale cross-linguistic work on the quantitative properties of languages has been largely restricted to what can be gleaned from minimally annotated text. For example, the Zipfian distribution of word frequencies has proven to be a robust universal of human language which continues to inspire research (e.g., Zipf, 1936; Mandelbrot, 1953; Plantadosi, 2014). However, comparative work on quantitative properties of syntax has only been done for small numbers of languages (e.g., Park & Levy, 2009) or small amounts of data (e.g., Bossong, 1998). Using the recent outpouring of standardized dependency-parsed corpora in many languages (e.g., Nivre et al., 2007; Zeman et al., 2012; McDonald et al., 2013), we have developed and evaluated a number of measures of morphological informativity, variability in dependency length, and variability in word order defined over corpora of parse trees. Here, we use these measure to characterize two key aspects of quantitative typology, word order freedom and dependency length, across a large number of languages.

**Word order freedom:** A long-standing generalization holds that the degree of order variability correlates with the presence of case marking (e.g., Sapir, 1923; Siwers, 1998; McFadden, 2003). However it is not straightforward to compare languages on their word order variability using existing data and methods. For example, WALS classifies German as "lacking a dominant word order" while classifying Classical Latin as "SOV" (Dryer, 2013), but it is not clear whether German has more word order variability than Latin or even how to test this. We developed several simple measures of entropy over word order. Fig. 1 shows one such measure: the entropy over the order of subject and object within a clause. As predicted, there is a clear trend for languages with more case marking to have greater entropy than those with less.

**Dependency length:** Dependency length has long been thought to be a crucial way in which human cognitive capacity interacts with syntax and has been the focus of much attention in linguistics and cognitive science (e.g., Gibson, 1998). In 20 languages, we examined the frequency distributions of dependency lengths and found that all follow a roughly power law distribution, with 1-word dependencies most common and a fat tail of longer dependency lengths (Fig. 2).

**Conclusion:** Besides the work presented here on word order freedom and dependency length, we believe that the measures and methods developed here can be used to quantitative answer many long-standing questions about cross-linguistic syntactic phenomena.

![Figure 1. Entropy of order of subject and object across 23 languages.](image1)

![Figure 2. Log-log plots of frequency distributions of dependency lengths across 20 languages.](image2)
THE ROLE OF GLOBAL SYNTACTIC CONTEXT IN STRUCTURAL PRIMING: EVIDENCE FROM GERMAN

Gunnar Jacob (Potsdam Research Institute for Multilingualism, University of Potsdam), Kalliopi Katsika, Alina Kholodova, Neiloufar Family, & Shanley Allen (University of Kaiserslautern)

gujacob@uni-potsdam.de

A key question in sentence production research is what information the formulator makes reference to when choosing between two alternative structures, such as between a prepositional object (PO) and a double object (DO). With regard to this question, Branigan, Pickering, McLean & Stewart (2006) distinguished between a local account, which assumes that the formulator is only influenced by properties of the elements which are part of the structure (e.g. in the case of a PO or DO, properties of the verb and the two noun phrases), and a global account, which assumes that the processor is influenced by structural properties of the entire sentence (e.g. also whether the PO or DO is part of a main clause or a subordinate clause). Theoretical accounts of formulation (e.g. Pickering & Branigan, 1998) explain formulation through the activation of combinatorial nodes, implicitly assuming a local account. Previous experimental studies (Branigan et al. 2006) failed to show an influence of context on formulation, a finding which supports a local account.

In a German structural priming experiment with sentence completion, we investigated whether structural priming effects are facilitated if prime and target share the same global structural context. Importantly, in German, global structural properties (e.g. whether a clause is a main clause or a subordinate clause) have direct consequences for constituent order within a clause. Eighty native speakers of German completed target fragments which could be completed by either a PO or a DO; these were preceded by prime sentences. We experimentally manipulated prime type (PO prime vs. DO prime) and prime/target similarity with regard to global structural context (same clause type in prime and target vs. different clause types in prime and target).

Crucially, in addition to a main effect of prime type (with more PO completions after PO primes irrespective of global structural context), the results also showed a significant interaction, with larger priming effects if prime and target were similar with regard to global structural context. This effect is difficult to explain by the local account.

We conclude that, at least for a language such as German, formulation is not only affected by properties of the elements within the structure to be computed, but also by the global structural context surrounding this structure.

References


RESOLVING LEXICAL AMBIGUITY USING SUB-PHONEMIC DURATION CUES: DID YOU HEAR “PLACE KIN” OR “PLAY SKIN”? 

Yoonjeong Lee (University of Southern California), Elsi Kaiser (University of Southern California) & Louis Goldstein (University of Southern California) 
yoonjeol@usc.edu

Lexical ambiguity arises with ‘open juncture’ sequences such as ice cream/ice scream. Prior research on word recognition shows that listeners can use sub-phonemic/allophonic cues (e.g., McMurray, Tanenhaus & Aslin, 2009). However, most prior work on ambiguous open juncture sequences has tested situations with clear prosodic boundary cues. Little is known about whether, in connected speech contexts without strong prosodic boundary cues, listeners exploit local sub-phonemic cues such as voice onset time (VOT) to segment these ambiguous sequences. If there are two competing segmentation options, which do not differ in syntactic or semantic plausibility, do sub-phonemic cues guide lexical ambiguity resolution? This question can help us to better understand how bottom-up acoustic cues interface with higher-level processes such as segmentation and word recognition in a situation where other constraints (e.g., phonotactics, syntax, semantics) do not rule out either segmentation option. We used mouse-tracking to examine segmentation of English internal open juncture sequences embedding /s/-stop sequences (e.g., base pam vs. bay spam, trace tab vs. tray stab, place kin vs. play skin). These sequences differ in the degree of aspiration of the stop (Lehiste, 1960): The voiceless stop in a cluster (e.g., /k/ in skin) has a shorter VOT (less aspiration) than a word-initial voiceless stop (e.g., /k/ in kin). We tested if listeners use these VOT cues to segment ambiguous sequences.

Elicitation: Four speakers produced ambiguous /s/-stop sequences in carrier sentences which placed them in phrase-internal unaccented positions. The ambiguous sequences (e.g., place kin vs. play skin) were extracted and spliced: VOTs for the stops in /s/##/C/ syllables (‘place # kin’) were replaced (i) by VOTs of another /s/##/C/ syllable (identity-spliced), (ii) by VOTs of #/sC/ syllables (cross-spliced). VOTs for the stops in #/sC/ syllables (‘play # skin’) were replaced (i) by VOTs of another #/sC/ syllable (identity-spliced) or (ii) by VOTs of /s/##/C/ syllables (cross-spliced). (Before splicing, we checked that the VOTs of the voiceless stops pattern as expected.) Below, we use ‘intended target’ to refer to the original word (e.g., kin in the /s/##/C/ condition and skin in the #/sC/ condition), and ‘competitor’ to refer to the other member of each pair (e.g., skin if kin is target, kin if skin is target).

Experiment: Participants (n=20) heard two-word sequences (e.g., place kin), while viewing a display with the second word and its competitor in the top L/R corners (positions counterbalanced; e.g., KIN (intended target) and SKIN (competitor)). Semantic associations between words were minimized, and frequency was controlled. The task was to click with the mouse on the second word they heard. Mouse clicks and trajectories were recorded.

Results & Conclusion: Although mouse trajectories did not differ significantly, click responses show significant effects of VOT sensitivity: When asked to click on what they heard, listeners chose the intended target over the competitor significantly more in the identity-spliced condition than the cross-spliced condition—i.e., there was interference from the competitor in the cross-spliced condition. Thus, VOT cues can bias listeners’ interpretation of the possible location of a word boundary between /s/ and the following stop, even when other cues in the acoustic signal (e.g., duration of /s/) point to the alternative segmentation. In sum, our study suggests that segmentation and lexical access are highly attuned to bottom-up phonetic cues; this kind of information may be stored as abstract knowledge in the listeners’ lexicon (e.g., gestural representations, Browman & Goldstein, 1986).

THE RECOGNITION OF ASSIMILATED FORMS DURING REAL-TIME COMPREHENSION: EFFECTS OF PHONOLOGICAL CONTEXT AND PRIMING
Mercedeh Mohaghegh & Craig Chambers (University of Toronto)
mercedeh.mohaghegh@mail.utoronto.ca

The process of spoken word recognition involves the incremental mapping of speech sounds to lexical candidates. However, variability in the production of speech sounds can result in ambiguity at the level of lexical selection. One common source of such variability involves phonological processes such as place assimilation. For example, in English the place of articulation of a coronal nasal or stop consonant can assimilate to a following labial sound such as [b], thus ‘phone box’ is perceptually similar to ‘foam box’ (/n/ → [m]). Some studies suggest the acoustics of the assimilated sound partially match both an unassimilated /m/ and /n/, providing cues to help compensate for (i.e., “undo”) assimilation (e.g., Gow, 2002). Further, studies have also suggested listeners can compensate for the effects of assimilation when a triggering context is present (e.g., when the sound precedes a labial consonant; Gow & McMurray, 2007). The current study focuses on the nature of the factors involved in compensation for assimilation during lexical access for both nasal and stop consonants.

A Visual World methodology was used in which English native speakers viewed displays containing five buttons that were each labelled with a printed word. On critical trials, the labels consisted of: a “prime” word (lean), a target word (dine), a competitor for the target differing only in place of articulation of the final sound (dime), and two unrelated distractor words. Each display was accompanied by two recorded sentences: a prime instruction (e.g., Click on the ‘lean’ button) and a subsequent target instruction (e.g., Now click on the ‘dine’ button), where the final consonants in prime and target words were each varied to be either assimilated or unassimilated. Critically, the initial sound of ‘button’ creates a phonological context where assimilation is viable, entailing potential ambiguity between targets and competitors. Prime sentences were included to test if compensation for assimilation in the target sentences is greater when assimilated forms (vs. unassimilated forms) had just been heard in the previous sentence. Expt 1 examined nasal consonants (as exemplified above) and Expt 2 examined stop consonants (e.g., ‘cat’ button, which, when assimilated, sounds like ‘cap’ button). For both experiments we used a comparatively large stimulus set (24 item sets) that was balanced for lexical frequency and orthographic complexity.

When target words were assimilated, listeners showed a tendency to click on the competitor button, suggesting they did not recognize that the surface acoustics might have arisen from assimilation. Moreover, this tendency was greater for nasals than for stops. We then examined online comprehension patterns for trials where participants selected the target button. Eye movement data from these trials showed a priming effect in Expt 1 (nasals) such that hearing a previously-assimilated form increased early saccades to target words (e.g., dine) within a time window focused on the final consonant and the following /b/ (compared to when the prime sentence was unassimilated). Thus having just heard an assimilated nasal increases listeners’ implicit expectation that an [m] they are hearing in an unfolding word is actually an assimilated /n/. However, this effect was not observed in Expt 2 (stops).

Overall the results suggest that compensation for assimilation may be less automatic than previously thought. Yet when compensation does occur, the effects seem to be more strongly tied to recent processing experiences (cf. priming) for nasals compared to stops. This may reflect the idea that contextual factors influence compensation for nasal assimilation at earlier stages of processing (Mitterer, 2011). Ongoing studies using the same task will provide additional baseline data on listeners’ identification patterns when the relevant phonological context (the /b/ in button) is no longer present.
Allophones are proposed to be linguistically indistinguishable entities on an abstract level, subsumed by their representation as a single phoneme during preparation phases preceding articulation of an utterance. This theory is investigated here following demonstration by Damian and Dumay (2009) of a behavioural priming effect in naming latencies to pictures using phrases in which a phoneme is repeated (for example, green goat). These findings have been interpreted as phonemic priming, and served as evidence for the role of phonemes in speech production. An alternative explanation, however, is with reference to the lower-level fact that very similar articulatory actions are required in each case to produce a very similar acoustic signal.

The present study aimed to dissociate these two alternatives. If Damian and Dumay’s findings are indeed due to abstract phonemic priming, we expect this to persist in cases where the initial segments represent the same phoneme but do not share an articulatory identity. Allophonic variants are exactly such cases, e.g. English velar stops /k/ and /g/ are palatalised before front vowels and velarized before back vowels. As a result, the initial consonants in the words kitty and coffee, while representing the same phoneme /k/, are articulatorily and acoustically distinct.

23 participants used Possessor+Noun phrases (e.g. Kitty’s kettle) to describe target pictures of 4 different characters (e.g. Kitty, Sidney) who were displayed holding common objects. Naming latencies from the onset of stimulus display were the dependent measure and error rates were examined to confirm compliance with the task. 200 critical trials, involving 40 different combinations of character and object, were completed by each participant, with each trial representing one of the following conditions.

+Phoneme, +Articulation The onset phones match at the level of phonology and also at the level of allophony, where applicable (Kitty’s kettle, Sidney’s celery)
+Phoneme, –Articulation The onset phones are two different allophones of the same phoneme (Kitty’s coffee, Kumar’s kettle)
–Phoneme, –Articulation The onset phones are phonemically unrelated and acoustically dissimilar (Kitty’s celery, Sidney’s kettle)

The presence of two different allophones of a single phoneme in a noun phrase (+Phoneme, –Articulation) produces behavioural priming to an even greater extent than the presence of two instances of a phoneme that has only a single realisation (+Phoneme, –Articulation: 998 ms vs +Phoneme, +Articulation: 1035 ms). Planned comparisons supported by a linear mixed effects model showed that both were facilitated compared to the control condition (–Phoneme, –Articulation: 1069 ms (p < 0.001 and p = 0.002 respectively), and also differed significantly from each other (p = 0.003).

Some further investigation remains to be done, but the facilitation demonstrated in the case of differing allophones (+Phoneme, –Articulation) provides initial evidence supporting the suggestion that the priming effect of interest is not simply due to acoustic or articulatory characteristics of the speech signal; rather, the psychological status of the phoneme is an important component.

References

Previous work has suggested that native speakers of Mandarin Chinese use syllables (rather than phonemes) as a basic unit of speech planning (e.g. O'Seaghdha et al., 2010). One line of evidence for this conclusion comes from the implicit priming paradigm, in which participants memorize spoken responses to a set of prompts. These responses either share some initial phonetic material, or do not. L1 English speakers produce responses more quickly when they share as little as an initial phoneme, suggesting that English speakers “pre-plan” the initial phoneme when it is shared for all responses. This, in turn, indicates that phonemes are a basic unit of planning in English. When the same task is carried out in Mandarin by L1 Mandarin speakers, reaction times are faster when responses share an initial syllable, but not when only an initial phoneme is shared, indicating a syllable- rather than phoneme-based production strategy for Mandarin (O'Seaghdha et al., 2010).

This cross-linguistic difference raises the question of which strategy bilinguals use when their L1 and L2 have different planning units. We explore this question using L2 English production by Mandarin-English bilinguals (where the L1 is Mandarin) as a test case. Bilingual speakers’ planning strategy may be completely determined by the dominant language, which would predict use of a syllable-based strategy in English for L1 Mandarin speakers. Alternatively, speakers may be able to switch strategies based on the task language, which would predict the use of a phoneme-based strategy in English. Verdonschot et al. (2013) found support for this hypothesis based on phoneme priming in a reading task. However the use of orthographic stimuli leaves open the possibility that the priming was due to faster visual processing of the shared initial letter.

In our experiment, speakers memorized responses (presented auditorily in order to avoid direct influence of orthography) to sets of three picture prompts and then produced these responses when the pictures were presented. Homogeneous sets of responses either shared an initial phoneme (homogeneous-phoneme condition), or an initial syllable (homogeneous-syllable condition). Heterogeneous sets were composed of recombined items from the homogeneous-phoneme and -syllable conditions such that they shared no initial phonetic material. Participants included late Mandarin-English bilinguals (average age of English acquisition of 10 yrs) and English monolingual controls. Table 1 shows average response times for each group and each condition (132 trials/subject/condition) with standard errors in parentheses. These results are part of a larger project that investigates both speech perception and speech production strategies in L1 Mandarin speakers, Mandarin heritage speakers, and L2 Mandarin speakers (L1 English).

As expected, English monolinguals show faster reaction times when an initial phoneme is shared, indicating a phoneme-based production strategy. This effect is larger when a whole syllable is shared, probably due to additional overlap in phonetic material. The Mandarin-English bilinguals follow the same pattern as the English monolinguals, supporting the hypothesis that bilinguals can adopt the typical production strategy of the task language.

Table 1: Preliminary data from five (out of a total of twenty) participants in each group

<table>
<thead>
<tr>
<th>Group</th>
<th>homogeneous phoneme</th>
<th>heterogeneous phoneme</th>
<th>difference</th>
<th>homogeneous syllable</th>
<th>heterogeneous syllable</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 English</td>
<td>749 ms (5)</td>
<td>755 ms (5)</td>
<td>6ms</td>
<td>785 ms (6)</td>
<td>811 ms (6)</td>
<td>26ms</td>
</tr>
<tr>
<td>L1 Mandarin</td>
<td>776 ms (10)</td>
<td>790 ms (6)</td>
<td>14ms</td>
<td>803 ms (8)</td>
<td>872 ms (10)</td>
<td>69ms</td>
</tr>
</tbody>
</table>

References:
Skilled readers extract different types of linguistic information (orthographical, phonological, morphological, semantic) from the parafovea during reading (Schotter, Angele & Rayner, 2012). Although these preview effects have been shown to drive efficient reading in skilled readers, it remains unclear how much and what kinds of information are used by beginning readers. According to the dual-route theory (Grainger & Ziegler, 2011), reliance on phonology-driven sublexical processes should be increasingly replaced by the use of faster, orthography-driven lexical processes as readers become more skilled. The priming of a target word (e.g., *nose*) with a pseudohomophone (e.g., *noze*) is typically used as a marker effect (PsH) for phonological processing. Similarly, the priming of a target word with a transposed letter nonword (e.g., *nsoe*) can be used as a marker effect (TL) for orthographic processing. Lexical decision experiments have shown that PsH effects do indeed decrease with reading age, while TL effects tend to be stable or to increase (Grainger, Lété, Bertrand, Dufau & Ziegler, 2012). It is not, however, clear whether these effects can also be found in lexical processing while reading connected text. To investigate this we used the boundary eye-tracking method to compare parafoveal PsH and TL effects in children and adults during silent sentence reading.

Stimulus sentences were presented in German with embedded target words which were initially presented as a misspelled preview until the readers’ eyes crossed a boundary directly in front of the target. PsH effects were measured by comparing fixation durations on the target word (e.g., *clue*) following a pseudohomophone preview (e.g., *clew*) in contrast to a control preview (e.g., *clom*). Likewise, TL effects were measured by comparing fixation durations on a target (e.g., *band*) following a transposed letter preview (e.g., *bnad*) in contrast to a control preview (e.g., *bcod*). Our results show strong PsH preview benefits for children across fixation duration measures but none for adults. We also found TL preview benefits for both children and adults. For adults, the effect was evident in measures of single and first fixation duration while it was observable in measures of single fixation and gaze duration for children.

We draw three main conclusions from our study. First, our results clearly show that the boundary method, which has so far only been used with adults, can also be used to measure parafoveal processes in children during silent reading. Second, we were able to replicate the developmental differences in PsH and TL effects which are typically found in lexical decision experiments. Finally, the pattern of effects indicates that the central predictions of the dual-route model also hold for children’s and adults’ parafoveal processing during silent reading in German. However, the TL effects for children shown in this and other studies of shallow orthographies further suggest that the expression and development of the dual route processes may be mediated by the orthographic depth of the language being learned.

Current theories of language production debate the extent to which language form is selected to facilitate comprehension, and the extent to which it is constrained by the production processing architecture. We examined this question in the domain of prosody. Speakers frequently use prosody to mark information status. For example, in the phrase “The panda blinks”, the pronunciation of “panda” is likely to be shorter when it is given (previously mentioned) than when it is new. The standard explanation is that prosody encodes the information structure, as a part of the intended meaning. Given information is topical, while new or contrastive information is focused, and needs to be marked with an accent. Yet this effect may also be influenced by the fluency of speech production processes. Given information is primed, which can speed production and promote reduced pronunciations, which sound less accented.

Our production experiment examined the role of each of these explanations. 36 speakers described images of animals performing pairs of actions (spin, expand, blink, shrink), in four conditions. The target is the second instruction:

1) Given: The panda spins. The panda blinks.
2) New: The frog spins. The panda blinks.
3) Compound: The panda and the frog spin. The panda blinks.
4) All: All the animals spin. The panda blinks.

The “Given” condition was compared with three conditions in which the target was focused, and thus should receive prominent prosodic marking (New, Compound, and All) under the standard account. By contrast, the fluency account predicted shorter/reduced pronunciations in both the Given and Compound conditions: when the target name had been mentioned, it should enhance fluency and support the production of short, unaccented forms.

As Figure 1 shows, the latency to begin speaking was shorter in the given (compared to new) condition, and in the compound (compared to all) condition. Thus, for each pair, planning was easier when the target had been mentioned. In addition, multiple-animal conditions (compound and all) had longer latencies. Target durations were also shorter for given (vs. new) and compound (vs. all). There was no effect of linguistic focus on the durations of the target word (panda), in that the compound condition patterned with given, rather than either new or all. The pitch data followed a similar pattern, although the difference between conditions was marginal.

These findings highlight the effects of utterance planning and effects on prosodic prominence. This is consistent with evidence that speakers reduce on the basis of their own experience, and that fluency promotes reduced pronunciations. Contrastive contexts can also elicit focus prosody, but here this effect was not observed here. Instead, we observed consistent effects of prior mention, which led to reduced prosody, regardless of whether the target was focused. We conclude the production fluency has a strong influence on prosodic form, and correlations between given status and fluency may contribute to information structure patterns.

THE TIME COURSE OF SEMANTIC CONSTRAINTS IN PRONOUN RESOLUTION
Israel de la Fuente & Barbara Hemforth (Université Paris-Diderot)
(israel.delafuente, barbara.hemforth)@linguist.univ-paris-diderot.fr

Research on discourse processing has shown that language comprehension is, to a significant extent, driven by expectations and that we use (extra)linguistic information incrementally to anticipate upcoming discourse. Following this assumption, we investigate how focus-sensitive particles (even and only) interact with discourse connectives (because and although) in the construction of discourse expectations, with specific consequences to pronoun interpretation in French, in sentences like (1).

(1) a. Pierre a interrompu Paul hier soir parce qu'il a/bien qu'il ait l'habitude de…
   'Peter interrupted Paul last night because/although he has the habit of…'
   b. Même Pierre a interrompu Paul hier soir parce qu'il a/bien qu'il ait l'habitude de…
   'Even Peter interrupted Paul last night because/although he has the habit of…'
   c. Seul Pierre a interrompu Paul hier soirée parce qu'il a/bien qu'il ait l'habitude de…
   'Only Peter interrupted Paul last night because/although he has the habit of…'

Focus-sensitive particles represent an ideal test-case here because of their multifactorial nature: not only do they increase the salience of the associated constituent, which is prosodically marked by the highest pitch accent in the sentence, but they additionally carry an important intrinsic semantic load (presuppositions, assertions, implicatures), which, in combination with certain discourse connectives, can create biases for a given referent over another (e.g. Filik et al., 2009; Moxey & Sanford, 1999; Koornneef & Sanders, 2013).

In particular, we predict that final interpretations will depend on expectations that are triggered by missing or unspecified causal content, in combination with a general processing strategy to avoid accommodation (see Bott & Solstad, 2014 for a similar proposal about implicit causality verbs). In (1b), a preference for the subject should arise in the concessive condition as a consequence of the expectation for a missing reason for the unlikeliness of Pierre, who is the least likely person to interrupt Paul, doing so. In (1c), this preference should arise in the causal condition as a result of the expectation for an explanation for the exhaustiveness of the entity in its scope, that is, why Pierre and nobody else interrupted Paul. If concessive subordinates can be taken as negative causals (e.g. König & Siemundi, 2000), we expect the opposite pattern, i.e. a preference for the object, as a result of the interaction of même+parce que and seul+bien que.

To investigate the interaction of these two factors in real-time language processing and in the construction of final discourse interpretations, we employed a task that combined Visual-World Eye-tracking and a continuation task: participants (n=31) were presented with images featuring four characters while they listened to sentence onsets, like those in (1), where the beginning of the subordinate clause can refer to either antecedent. They were first asked to provide an appropriate continuation to the sentence onsets orally and, at the end of each trial, to answer the question “Who has that habit?” by clicking on the image their continuation was about.

Our predictions were born out in the continuation data where a significant interaction between focus particles and connectives was found. For causal relations, participants showed a robust object antecedent interpretation with même and a subject antecedent interpretation with seul. This pattern was reversed for même (though not for seul) with concessive relations. Crucially, however, the eye-movement data revealed interesting patterns at the different time frames clearly separating the impact of different aspects of the semantics of the focus particles and connectives: at the onset of the connective, there is a clear effect of focus with lower proportions of looks to the subject antecedent in the two baseline conditions (1a) compared to the other four conditions with même and seul; at the pronoun (and “l’habitude de” phrase), there is a clear effect of discourse connective, with a stronger preference for the subject antecedent in the concessive conditions compared to the causal conditions; and, finally, during the continuation, we found exactly the same significant interaction between the two factors as for the continuations.
INFORMATION DENSITY AND THE USE OF DISCOURSE CUES

Fatemeh Torabi Asr and Vera Demberg
MMCI Cluster of Excellence, Saarland University, Germany

The Uniform Information Density (UID) hypothesis [3] holds that speakers generally formulate their messages such that a uniform rate of information transmission is achieved. UID hence predicts that optional linguistic elements are used when the information they convey is unpredictable, and that they may be omitted when they convey little new information [1,4]. Discourse relations can be signaled by discourse connectives (e.g., “because” for a reason relation), which are one kind of optional linguistic element. Asr and Demberg [1] showed that discourse connectors are more likely to be present in discourse relations which are less expected based on general cognitive biases, providing support to the UID hypothesis. However, it has not yet been shown that local linguistic context also affects the distribution of discourse connectors.

We here argue, based on a corpus study using the Penn Discourse Treebank [5], that certain types of linguistic features, such as negation words, can also signal discourse relations (e.g., *chosen alternative* [6]), such that the discourse connective signaling the same relation would be partially redundant, and hence more likely to be omitted. We perform a normalized point-wise mutual information (NPMI) analysis [2] of the explicit and implicit discourse relations in the PDTB, with negation in the first argument of the relations as a cue:

\[
NPMI(relation \mid cue) = \frac{\log p(relation)p(cue)}{\log p(relation,cue)} - 1
\]

Positive NPMI values mean that a discourse relation is more likely to be present following negation, while NPMI is negative if a discourse relation is less likely given the negation cue. As Figure 1 illustrates, the analysis provides evidence that 1) negation in Arg1 of a discourse relation changes the distribution of possible relation senses (dark bars), 2) negation has an even larger effect on implicitness of the relation (light bars). Previous studies on connectives emphasize their facilitative role in processing discourse relations. Based on our corpus analysis, we propose an explanation for the usage of connectives by language producers: Connectives are more likely to be used when the discourse relation is not predictable, and omitted when the discourse relation is expected due to a linguistic cue in the local context. These findings support the UID hypothesis at the level of discourse relations.

![Figure 1: NPMI calculated for frequent relations (dark green) and their implicit occurrences separately (light orange). Relations written in bold face are positively marked by negation in Arg1, whereas, the other relation senses are found to be negatively correlated with the negation cue.](image)

MUCH OF FLIES AND INSECTS. CROSS-LANGUAGE ACTIVATION AND INHIBITION IN INTERLINGUAL HOMOGRAPHS COMPREHENSION.

Joanna Durlik, Jakub Szewczyk, Zofia Wodniecka (Jagiellonian University in Kraków)
joanna.durlik@uj.edu.pl

Interlingual homographs (IHs) are words that share orthography but differ in meaning across languages (e.g. an English-Polish homograph “much” means “flies” in Polish). Due to their ambiguity, they are frequently used to investigate how (non)selective is language activation in bilingual comprehension.

In the present study unbalanced Polish-English bilinguals completed a task containing Polish-English IHs, which was based on Macizo, Bajo and Martin's (2010) Spanish-English task. Participants were asked to decide whether presented pairs of English words were semantically related. The task was composed of 40 two-pair blocks. In 20 blocks, the first pair consisted of an IH matched with an English word related to IH's Polish meaning but unrelated to its English meaning (e.g. “much – insects”). In the remaining 20 blocks, the first pair consisted of unrelated English words with no meaning in Polish (e.g. “oily – insects”). Manipulation in the second pair was the following: in 20 blocks an exact English translation of the Polish IH’s meaning was paired with a related word (e.g. “flies – buzz”), whereas in the remaining 20 blocks, the translation was replaced with a control word related to the second word in the second pair (e.g. “bee – buzz”). Types of the first and the second pairs were fully crossed, yielding four experimental conditions: IH-Translation, IH-Control, Control-Translation, Control-Control.

Macizo’s et al. (2010) found that RTs to the second pair in IH-Translation condition were significantly longer than in the Control-Translation condition (p < 0.001), whereas RTs to the second pair in IH-Control and Control-Control conditions did not differ significantly. These effects indicated that irrelevant meanings of IHs were initially co-activated and then inhibited.

In the present study, surprisingly, in the second pair both the IH-Translation and IH-Control conditions yielded longer RTs, relative to the Control-Translation and Control-Control conditions (ps < 0.001). These results lead to two major conclusions. Firstly, they replicate evidence for initial co-activation of both L1 and L2 IHs’ meanings, which supports Macizo’s et al. (2010) findings. Secondly, the present results suggest that L1 inhibition affects not only particular concepts linked to irrelevant IHs' meanings, but operates more globally and may hamper access to entire semantic categories which the concepts are related to. For example, when an English-Polish homograph “much” elicits a concept of a FLY (relevant for the Polish meaning of the word “much”), other related concepts may also get inhibited (e.g. concepts of a BEE or BUZZING). Therefore, the present study suggests that the scope of inhibition of contextually irrelevant L1 in IHs processing may be broader than previously believed, and that inhibition targets the conceptual, rather than lexical level of representation.
NO SQUIRREL LIKES COLLECTING NUTS.
AN ERP STUDY ON QUANTIFICATION, PREDICTION, AND CLOZE PROBABILITY.
Dominik Freunberger & Dietmar Roehm (University of Salzburg)
dominik.freunberger@sbg.ac.at

When is the information a quantifier provides available in online sentence processing? And how does a sentence-initial negative quantifier influence the processing of words that are strongly predictable in the non-negated/affirmative version of the sentence?

We used 60 German affirmative sentences where the sentence-final noun was either highly predictable (1A) or unpredictable but still plausible (1B) according to a cloze probability (CP) pre-test. To test the influence of quantification, an initial negative quantifier was added to those 60 sentences (2A/B), resulting in a total of 120. We recorded the EEG of 23 German native speakers while they were silently reading the sentences word by word. After each sentence, participants were asked to judge the sentences’ acceptability. ERPs were calculated for the sentence-final nouns (see table).

<table>
<thead>
<tr>
<th>Example (CW underlined)</th>
<th>CP</th>
<th>ACC (%)</th>
<th>RT (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A Eichhörnchen sammeln gerne Nüsse.</td>
<td>.93</td>
<td>99.28</td>
<td>553.55</td>
</tr>
<tr>
<td>1B Vögel fangen häufig Ratten.</td>
<td>.03</td>
<td>88.64</td>
<td>699.12</td>
</tr>
<tr>
<td>2A Kein Eichhörnchen sammelt gerne Nüsse.</td>
<td>.15</td>
<td>2.62</td>
<td>595.10</td>
</tr>
<tr>
<td>2B Kein Vogel fängt häufig Ratten.</td>
<td>.01</td>
<td>11.52</td>
<td>816.78</td>
</tr>
</tbody>
</table>

As in former studies, high-CP words (1A) led to a posterior P300 (see figure), which has been related to the integration of correctly predicted words [1, see also 2]. Unexpected, low CP words in the affirmative (1B) and in the negated sentences (2B) engendered a biphasic N400-P600 pattern [e.g. 3]. Strikingly, despite the low CP (.15), critical words from 2A led to a comparable P300 as high CP (.93) words from 1A. Possible interpretations of these results are: 1) The initial negative quantifier was not processed immediately [cf. 4], because this would have led to an N400 in 2A (in accordance with the low CP and acceptability). 2) The P300 and N400 are more sensitive to semantic than lexical properties (e.g. relatedness vs. CP), since significantly different CPs led to comparable ERPs, when semantic properties like relatedness were identical between affirmative and negated sentences. This implies 3) that CP might be a suboptimal offline measure for investigating prediction, because it merely quantifies the likelihood of a lexical item within a context but does not include semantic properties; this can, like in our study, lead to a low probability of a word in a specific context although high semantic relatedness to its context is given (2A).

INCREMENTAL PROCESSING OF QUANTIFICATIONAL RESTRICTION – EVIDENCE FROM CROSS-MODAL PROBE RECOGNITION

Petra Augurzky, Oliver Bott, Wolfgang Sternefeld & Rolf Ulrich (Universität Tübingen)
petra.augurzky@uni-tuebingen.de

Background: While theories on sentence comprehension generally agree that the structural integration of a newly incoming word takes place incrementally, it is still an open issue whether compositional semantic interpretation proceeds in a comparable fashion. For instance, though previous studies have shown immediate effects of quantifier interpretation (e.g. Wijnen & Kaan, 2006), recent evidence suggests that at least some aspects of quantificational meaning are evaluated in a delayed manner (e.g. Bott & Schlotterbeck, 2013).

Methods: We examined pictorial context effects on the processing of questions involving quantificational restriction ((1a/b), Fig.1). As a complete interpretation (answer to the question) is possible locally on the colour adjective, a strong version of incrementality predicts a truth evaluation directly on this position, irrespective of whether a following restrictive cue will change the locally assigned value. Local and global (clause-final) evaluations differ for some contexts, thus leading to potentially required meaning updates. For assessing whether truth values are evaluated locally, we ran a dual task: After contexts were shown, sentences were presented via RSVP (500ms/word). Participants made a judgment about a tone that was heard 400ms after adjective onset (Did the tone appear on the left or the right ear?), and a clause-final truth value judgment about the picture-sentence pair (Is the sentence true w.r.t. the contents of the picture?). Participants (n=64) had to evaluate both tasks as quickly as possible. If tones were presented to the left (50%), participants had to react with their left index finger, otherwise they had to react with their right index finger. Buttons (true/false) for the truth evaluation were counterbalanced across participants. We expected response preparation (facilitation) when truth and tone evaluations triggered the same motor response.

Hypotheses: H1. A strong version of incrementality predicts faster reaction times for the tone evaluation in all contexts when the answer button for the tone task corresponds to the button for the locally calculated meaning. H2. If quantificational truth evaluation is generally delayed, no mid-sentence facilitation will be observed for the tone task, but clause-final decisions should correspond to global judgments. H3. If local judgments are only made in the absence of a risk of reanalysis, local facilitation for B and C is expected, as the final semantic value is already available on the adjective. No such facilitation is expected for A and D, as potentially following information might cause a shift in the locally assigned truth value.

Results: Our findings speak in favour of the incrementality hypothesis H1. Reaction times for the sound task were faster in case of a match between the button for the locally expected truth value and the button for the tone task (Effect of Identity: F(1,63) = 5.17; p<.05). Moreover, an effect of Context (F(3,189); p<.001) showed a facilitation of Context B as opposed to all other contexts (all p’s<.001). This additional facilitation for B cannot be explained straightforwardly by any of the mentioned hypotheses. Contrary to the predictions of H3, there was no interaction between Context and Identity (p>.2). For the truth evaluation, we did not find any costs for meaning updates, but a generally good performance (90% correct). In sum, our results indicate that the semantic evaluation of questions involving quantification proceeds rapidly and in an incremental manner. The present findings are the basis for future research employing language-related ERPs and the Lateralized Readiness Potential (LRP).

Figure 1

<table>
<thead>
<tr>
<th>(1) a. Sind alle Dreiecke blau, die im Kreis sind?</th>
<th>(1) b. Sind alle Dreiecke blau, die außerhalb des Kreises sind?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are all triangles blue, that are in the circle?</td>
<td>Are all triangles blue, that are out-of the circle?</td>
</tr>
<tr>
<td>Shift in truth value?</td>
<td>Global</td>
</tr>
<tr>
<td>A</td>
<td>True</td>
</tr>
<tr>
<td>B</td>
<td>True</td>
</tr>
<tr>
<td>C</td>
<td>False</td>
</tr>
<tr>
<td>D</td>
<td>False</td>
</tr>
</tbody>
</table>
LOOKING BEYOND ACCESSIBLE REFERENTS. EFFECTS OF DIFFERENT TYPES OF INDEFINITE NOUN PHRASES ON THE SUBSEQUENT DISCOURSE

Sofiana Chiriacescu (Universität zu Köln).
Sofiana.chiriacescu@uni-koeln.de

One of the fundamental questions underlying theories of language production concerns referent-tracking. A body of linguistic and psycholinguistic studies found out that several factors, such as prominent syntactic positions and different thematic roles, influence the frequency of re-mention and the rate of pronominalization of referents. In this paper, we focus on referents mentioned in non-prominent positions (as direct objects realized by indefinite noun phrases) in English and German and argue that different types of indefinite noun phrases affect the discourse structuring potential of their referents in different ways.

**Study1: The English data (Indefinite-*this* vs. indefinite-*a*)**

According to several studies (Prince 1981, Ionin 2006), English *this* can be used as an indefinite determiner alongside the simple indefinite article *a(n)*. The Experiment (Exp1) investigates whether referents introduced by *this* and *a(n)* differ in terms of frequency of subsequent mention and likelihood pronominalization. **Design.** We used a multi-sentence continuation task with no pronoun-prompt. Participants (*n*=20) read story fragments (e.g. (1), *n*=20) and were asked to add five logical and natural-sounding sentence continuations to each of the stories. All critical referents were constructed in direct object position and were realized as indefinite noun phrases. We only manipulated the morphological realization of the direct objects (2 conditions: *this*-condition and *a(n)*-condition). In light of previous findings, which showed that *this*-referents are more accessible than *a(n)*-referents (Gernsbacher & Shroyer 1989), we predict that *this*-referents will be: (i) more frequently picked up, and (ii) more likely to be pronominalized in the subsequent discourse, compared to *a(n)*-referents. **Results.** *This*-referents were picked up in the subsequent discourse more often than *a(n)*-referents (in 85% vs. 15% of the cases), but contrary to our predictions the anaphoric expressions used for both indefinite types were definite noun phrases.

**Study2: The German data (Indefinite-*so’n* vs. indefinite-*ein*)**

The German determiner *so’n* can be used in a similar way as English indefinite *this* (Chiriacescu 2010, 2011). Experiment 2 (Exp2) had the same design, but tested the discourse behavior of indefinite *so’n* compared to that of the simple indefinite headed by *ein(e)* (*a(n)*). Again, we manipulated only the type of indefinite noun phrase, which resulted in 2 conditions: *so’n*-condition and *ein(e)*-condition. **Our prediction is, that if the accessibility of *so’n*-referents is comparable to that of referents preceded by indefinite-*this*, then the results of the two experiments should be similar.** **Results:** Similar to the findings of Exp 1, *so’n*-referents were picked up more often in the ensuing discourse than the *ein(e)*-referents (in 80% vs. 17% of the cases), but did not show a preference for pronominalization. **Conclusions:** First, both indefinite *this* and indefinite *so’n* signal the referential persistence of their referents in the subsequent discourse. Second, the findings of both Exp1 and Exp2 underline the necessity to dissociate between frequency of subsequent mention and likelihood of pronominalization, as they do not point to the same type of accessibility of a referent (confirmation of recent studies on language production, e.g. Kehler, Kertz, Rohde & Elman 2008). Third, we argue that the different markers of indefinite noun phrases (i.e. *this* in English and *so’n* in German) are used as linguistic devices to give structure to the subsequent discourse rather than to signal the accessibility of their associated referents.

<table>
<thead>
<tr>
<th>this-condition</th>
<th>a(n)-condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yesterday evening was so warm that James decided to hang out with friends at the local coffee shop. On his way downtown, he saw <strong>this kid</strong> coming down the street.</td>
<td>Yesterday evening was so warm that James decided to hang out with friends at the local coffee shop. On his way downtown, he saw <strong>a kid</strong> coming down the street.</td>
</tr>
</tbody>
</table>

Table 1. Sample experimental item from Exp1 on German
HEMISPHERIC DIFFERENCES IN PROCESSING THE SCENARIO RELEVANCE OF WORDS IN DISCOURSE
Ross Metusalem, Marta Kutas, Thomas P. Urbach, & Jeffrey L. Elman (all UC San Diego) rmetusal@ucsd.edu

Discourse comprehension often entails inferring unstated elements of described scenarios or events through activation of relevant knowledge in long-term memory. Metusalem et al. [1] examined the degree to which unstated event knowledge elements are activated during incremental comprehension, specifically at points in a sentence at which those elements would constitute semantic anomalies. Participants’ EEG was recorded as they read short stories in which the final sentence contained one of three target word types: highly expected, anomalous but related to the described event, and anomalous but event-unrelated (see Figure 1). Metusalem et al. found a three-way split in the amplitude of the N400 event-related brain potential (ERP) component. Expected words elicited the smallest N400, event-unrelated the largest, and crucially event-related an N400 between these two extremes. This N400 pattern was argued to indicate that real-time activation of event knowledge occurs with at least partial independence from the immediate sentential context.

The present study addresses hemispheric contributions to the effect observed by Metusalem et al. While the left hemisphere (LH) has been argued to support expectations for likely upcoming words through semantic feature pre-activation [2], the right hemisphere (RH) is believed to activate concepts beyond those that would constitute expected continuations of the sentence [3]. We therefore hypothesized that RH activity might be driving much, if not all, of the difference in N400 amplitude between event-related and event-unrelated targets in Metusalem et al.’s previous study.

In the present experiment, Metusalem et al.’s stimuli were used, only now with target words presented to the right or left visual field (RVF/LVF) only. This visual half-field presentation provides a processing advantage to the hemisphere contralateral to the visual field of presentation. The results show that a three-way split in the N400 time window (and beyond) is found only with LVF/RH presentation, indicating that the difference in N400 amplitude between event-related and -unrelated anomalous words is driven largely by right hemisphere processes. This complements the findings of Federmeier and Kutas [2], by demonstrating that while the LH activates the semantic features of words likely to appear in the sentence, the RH more broadly activates concepts related to the described event.

| A huge blizzard ripped through town last night. My kids ended up getting the day off from school. They spent the whole day outside building a big ________ in the front yard. |
| Expected: snowman ____ |
| Event-related: jacket ______ |
| Event-unrelated: towel ________ |

Figure 1. Target ERPs at the midline parietal site for RVF/LH and LVF/RH presentation


Do temporal terms like ‘before’ and ‘after’ impact people’s ability to rapidly map incoming language onto real-world knowledge? People are slower and less accurate when reading sentences that start with ‘before’ compared to ‘after’ [1], reflecting people’s default expectation that narrated events occur in chronological order [2-3], flouted by sentence-initial ‘before’. However, extant research has examined comprehension of ‘before’ as subordinate conjunction that connect two events (‘Before [Event2], Event[1]’), not as preposition that relate events to fixed time-points (‘before dinner’), which is how children acquire ‘before’ first. The current study compared N400 truth-value effects [4] to sentences about known events, under explicit verification instructions or without (each N=30). Event-outcome knowledge may compete for activation when people comprehend sentence about before-states [5-6]. Words that are consistent with outcome knowledge may thus be ‘inappropriately’ facilitated and act as a ‘lure’ when they render sentences false, leading to a reduced N400.

**Methods:** Materials were 120 sentence quadruplets (conditions differed in temporal connective or critical word), selected so that true/false conditions were matched on cloze-value and truth-value pre-ratings (see figure). CWs were matched on length, frequency and semantic relatedness (LSA-SSV). Fillers were 200 sentences without temporal connectives. Sentences were presented word-by-word (300ms duration, 200ms blank). 64-channel EEG data was collected. Verification-instruction EEG data included trials with condition-consistent responses only (1-agree/5-disagree, +/-27 correct trials in all conditions).

**Results:** N400 analysis (350-450ms, 16 electrodes in 4 quadrants) showed a main N400 effect of truth-value (false>true, $F(1,55)=10.7$, $p<.005$), and a 2(distribution: anterior, posterior) by 2(connective: before, after) by 2(truth-value: true, false) 3-way interaction ($F(1,56)=4.6$, $p<.05$). At posterior sites, false-after sentences elicited larger N400s than false-before sentences ($M=-1.04$, $p<.05$), no difference was found for true sentences. No effects of instruction and no main effect of connective were observed. In the subsequent 600-800 ms window, false-before sentences elicited more positive ERPs than true-before sentences at posterior electrodes ($M=1.0$, $p<.05$), no such effect was observed for ‘after’.

**Conclusion:** Sentences about real-world events show reduced online effects of truth-value when they start with ‘before’ compared to ‘after’. ‘Before’ specifically impedes comprehension when language does not map onto world knowledge. Despite being judged equally false, words that are consistent with the event outcome are momentarily facilitated compared to words that are not. This inappropriate facilitation leads to an asymmetric impact of ‘before’ and ‘after’ when people map incoming language onto what they hold to be true.

---

HOW VERTICAL HAND MOVEMENTS IMPACT BRAIN ACTIVITY ELICITED BY LITERALLY AND METAPHORICALLY RELATED WORDS
Seana Coulson (UC San Diego), Megan Bardolph (UC San Diego)
scoulson@ucsd.edu

Introduction. Grounded theories of meaning implicate sensory motor regions in the coding of word meaning [1]. Previous work has shown that motor-related brain areas are involved in language processing (see [2] for review). However, prior research leaves open the stage or stages of processing at which motor activations occur (early vs. late), and whether their time course is similar in literal vs. metaphorical words – as predicted by metaphor comprehension models, such as conceptual metaphor theory. Here we compared event-related brain potentials (ERPs) elicited by words associated with upper and lower regions of vertical space, related either literally, (cloud/boot), or metaphorically, (victory/defeat), as healthy adults performed a concurrent motor task with either an upwards or a downwards trajectory.

Methods. Participants were 24 undergraduates (13 male) aged between 18 and 34 (mean=20) years. All were right-handed, had normal or corrected-to-normal vision, were fluent English speakers, and had no history of neurological or psychiatric disorders. Materials included 42 words literally associated with upper regions of space (Literal/High), 42 words literally associated with lower regions of space (Literal/Low), 42 words metaphorically associated with upper regions of space (Metaphor/High), 42 words metaphorically associated with lower regions of space (Metaphor/Low), and 84 filler words, associated with neither upper nor lower regions of space. Participants’ EEG was recorded as they read visually presented words while performing a one-handed motor task in which they were asked to move marbles from either a red tray to a green one (i.e. an upwards-directed movement), or from a green tray to a red one (a downwards-directed movement). Half of the participants performed the task with their right hand, and half performed the task with their left. The experiment was divided into 4 blocks, half of which involved moving marbles to the green tray, and half to the red. Participants saw each word twice, i.e. while moving marbles both upwards and downwards. Direction of initial movement was counterbalanced across participants. Words were presented for 500ms, followed by 1500ms of blank screen, and ERPs were time-locked to word onset.

Results. Congruency effects were observed both 200-300ms and 700-900ms post- word onset, as words incongruent with movement direction elicited slightly more positive ERPs. However, the early P2 congruity effects were only evident in words whose verticality was Literal (p<0.01), and absent from words whose verticality was Metaphorical (ns.). Congruity effects 700-900ms (p<0.05) did not significantly differ as a function of Literal/Metaphorical, suggesting movement direction similarly impacted the processing of both sorts of words. No congruity effects were observed 300-500ms post-onset, the interval associated with the N400, the ERP component most clearly related to retrieval from semantic memory.

Discussion. Here we show that brain activity elicited by the very same words differed as a function of the direction of hand movement. Moreover, these differences emerged much sooner for words literally related to the vertical dimension than for metaphorically related words. For literally – but not metaphorically – related words, movement direction impacted ERPs elicited 200-300ms, during a portion of the ERP waveform sensitive to both phonological and semantic factors. In later intervals, 700-900ms post-onset, both literally and metaphorically related words were similarly impacted by the movement manipulation, consistent with the claim that lexically induced motor system activation occurs during a relatively late stage of processing following the initial stages of semantic retrieval indexed by the N400 component. Results argue against strong embodied metaphor theories.

EFFECTS OF WORKING MEMORY LOAD ON THE PROCESSING OF NUMERALS
AND STANDARD SCALAR QUANTIFIERS

Laine Stranahan (Harvard University), Qingqing Wu (Harvard University), Manizeh Khan
(Harvard University), and Jesse Snedeker (Harvard University)
stranahan@fas.harvard.edu

Upper-bounded interpretations of scalar quantifiers like some (some but not all) take
longer in timed judgment studies (Bott & Noveck 2004) and appear later in processing
(Huang & Snedeker 2009) than lower-bounded ones (some and possibly all),
supporting the claim that the former are pragmatically derived from the latter. Working
memory load decreases rates of upper-bound-based interpretations of some in offline
tasks without affecting all, suggesting pragmatic but not semantic processing is impaired
by load (Marty, et al. 2013). We address two questions: (1) How does the effect of load
emerge in real time? and (2) Does load only affect pragmatics, or does it affect
processing more generally?

To investigate (1) we conducted an eye-tracking experiment (N=78) using the visual
world paradigm. We adopted stimuli from Huang & Snedeker (2011) but separated
scalars (some, all) from numerals (two, three) such that each participant heard only one
category of quantifier: In this paradigm, implicatures appear to be calculated most
robustly when only one conceptual dimension is targeted at a time (Grodner, et al.
2010). Participants in the memory load condition were first presented with a sequence
of (2 or 4) letters to memorize, then following a short story were instructed to point to
one of the characters on the screen—either the character with some but not all of the
objects (“point to the girl with some of the watermelons”) or the character with all of the
objects (“point to the girl with all of the watermelons”). Eye movements during the
instruction phase revealed if and when participants had computed an upper bound.
After pointing, participants typed the letter sequence in reverse order. (No-load trials
lacked the memorization and recall stages.) We found a significant increase in target
preference between the region before the quantifier and the region before the noun
phrase disambiguation in all trials (p < 3.34e⁻⁷) but not in some trials (p > 0.1). This
pattern was present in all three load conditions, suggesting that the implicature was not
calculated even when load was absent.

To better understand effects of load, we repeated the experiment (N=64) using the
numeral quantifiers two and three in place of some and all. In both two (p < 2e⁻¹⁶) and
three (p < 2e⁻¹⁰) conditions we saw a significant target preference, but unlike some and
all (p > 0.95), target preference was significantly modulated by load (p < 0.003).
Since we failed to observe implicature, our study offers no insights into (1), and future
work will explore several hypotheses for this failure. Meanwhile, our data do indicate
that working memory load can interfere with processing below the pragmatic level.

References: Bott, L., & Noveck, I. A. 2004. Some utterances are underinformative: The
onset and time course of scalar inferences. Journal of memory and language, 51(3),
“Some,” and possibly all, scalar inferences are not delayed: Evidence for immediate
Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics
2013. Interpreting numerals and scalar items under memory load. Lingua, 133,
152-163.
GAPS WITHIN FILLERS ARE NOT BETTER THAN GAPS WITHIN SUBJECTS
Grant Goodall (University of California, San Diego)
ggoodall@ucsd.edu

As is well known, gaps in filler-gap dependencies resist being located within a subject (the subject island effect):

(1) * Which animal will [several movies about _ ] be shown to the visitors?

It is often claimed also that these gaps improve greatly when the subject is itself a wh-filler (Chomsky 86, Kayne 84, Lasnik&Saito 92, Rizzi 06, Torrego 85; but see Müller 95):

(2) Which animal do you wonder [how many movies about _ ] will be shown to the visitors?

This is an astonishing claim from a processing perspective. (2) requires positing a filler (which animal) and then positing a second filler (how many movies about) while simultaneously resolving the dependency associated with the first filler (after about). The processing resources required would seem to be much greater than those needed for (1), thus predicting lower acceptability, so it is very surprising that acceptability would increase significantly. (2) is also surprising from the standpoint of syntax, since it seems to violate some well-motivated principles of grammatical architecture (Rizzi 04, Wexler&Culicover 80).

We propose here that despite the claims in the literature, placing gaps within fillers does not increase their acceptability. This misunderstanding has arisen because in many of the cited examples, pied-piping rather than preposition stranding is used, which allows for the gap to be parsed outside of the fronted filler. We demonstrate this by means of a formal acceptability experiment in English.

**Experiment.** 48 subjects rated sentences on a scale from 1 ("very bad") to 7 ("very good"). Materials were all w/h-questions which varied by the type of extraction (preposition-stranding vs. pied-piping), the location of the phrase containing the gap (fronted filler (SPEC/CP) vs. embedded clause vs. matrix clause) and the grammatical function of the phrase containing the gap (subject vs. object). Subjects saw 4 tokens of each condition (=48 experimental items), combined with 57 filler items. These were counterbalanced using a Latin Square design and pseudo-randomized, producing 12 lists. 12 additional lists had the reverse order. Results were converted to z-scores and are presented here (error bars represent SE):

With pied-piping (i.e. no preposition stranding), there is no contrast between gaps within subjects and gaps within objects, suggesting that in the subject case, speakers are able to posit the gap somewhere outside of it. With preposition stranding, on the other hand, the gap site is unambiguous, and a gap within a subject is clearly worse than within an object in the "embedded" and "matrix" cases, as expected. With gaps within a fronted filler (SPEC/CP), the improvement over gaps within a subject ("embedded" or "matrix") that is claimed in the literature does not materialize. This is welcome and reassuring news; going against basic principles of processing and/or grammar does not make the sentence more acceptable.
Some Cushitic languages are analyzed as having three values for gender: masculine (M), feminine (F) and “plural (P)”. Plural is a value given for nouns that trigger the same agreement form as third person plural even when they have singular reference. It is not entirely clear yet whether this third value, “plural”, is processed similarly to masculine and feminine. In this study, two picture-word interference experiments examined this issue in Konso, a Cushitic language spoken in Ethiopia.

The gender of nouns determines the assignment of definite (DEF) marking on nouns in Konso. Plural gender nouns take the definite suffix –siniʔ, and masculine and feminine gender take the definite suffix –siʔ. In Experiment 1, native Konso speakers produced nouns with a gender-marked definite suffix (e.g. innaa-siniʔP ‘boy-DEF[P]’ ‘the boy’ vs. karma-siʔM/F ‘lion-DEF[M/F]’ ‘the lion’) while ignoring simultaneously presented auditory distractor words, which have the same or different gender value as the targets. Naming latencies were significantly shorter in the gender-congruent condition than the incongruent condition.

Konso shows gender agreement in the subject inflection on the verb (feminine take -t, masculine take -ay, and plural gender take -n). In Experiment 2a, participants responded to target pictures by producing sentences (with an overt subject [name of a picture] and a verb [ʔi=akk-am- + gender marker] ‘shown’) while ignoring the distractor words. For instance, participants respond to a picture of a bone by producing lafta-siʔʔi=akk-am-t-3=show-PAS-3F-PF/* ‘The bone was shown’. In Konso, the overt subjects can be omitted and they can be understood from the gender agreement markers on the verb. In Experiment 2b, participants responded to target pictures by producing only verbs, ʔi=akk-am- + gender marker ‘shown’, while ignoring distractor words (e.g. ʔi=akk-am-t-i3=show-PAS-3F-PF/ ‘She was shown’ for a picture of a bone). In both Experiment 2a and 2b, compare to gender-incongruent condition, gender-congruent distractor words sped up the naming time of the target-pictures significantly and triggered fewer errors.

The fact that this gender congruency effect was also observed for the production of plural gender nouns provides evidence that plural gender is processed similarly to masculine and feminine. This supports analysis of singular nouns marked with the plural suffix as plural gender nouns. The results also demonstrate that the selection of gender-marked suffixes involves competition.


* 3 = third person, PAS = passive, PF = perfective
Author Index

Abashidze, Dato, 149
Abbot-Smith, Kirsten, 136
Ackerman, Lauren, 27, 75, 161, 228
Akahane-Yamada, Reiko, 142
Aleman Banon, Jose, 213
Allen, Shanley, 168, 247
Almeida, Diogo, 41, 90, 91
Almor, Amit, 68
Amenta, Simona, 113, 190, 202
Anderson, Andrew, 184
Apperly, Ian, 55, 58, 147
Apple, Aaron, 26
Arai, Manabu, 217, 218
Arnett, Nathan, 221
Arnold, Jennifer, 253
Arnon, Inbal, 24, 176
Atchley, Ruth Ann, 95
Atkinson, Emily, 219, 220
Augurzky, Petra, 258
Avrutin, Sergey, 207
Azevedo, Nancy, 95

Baayen, Harald, 97
Back, Elisa, 147
Bader, Markus, 193, 234
Bai, Chunhua, 166
Bak, Thomas H., 140
Balewski, Zuzanna, 23
Bardolph, Megan, 262
Barkley, Christopher, 242
Baumann, Peter, 161, 214
Begum, Hisma, 194
Benavides-Varela, Silvia, 22
Benjamin, Aaron, 61
Bernolet, Sarah, 178
Betts, Hannah, 229
Boon, Paul, 178
Bordag, Denis, 64–66
Bornkessel-Schlesewsky, Ina, 191
Böröndi, Anita, 131
Bosker, Hans Rutger, 87
Bott, Oliver, 258
Bottini, Roberto, 79
Bowers, Jeffrey, 198
Bradlow, Ann, 80, 251
Branigan, Holly, 131, 163
Braun, Bettina, 143
Brennan, Jonathan, 241
Brocher, Andreas, 69
Brouwer, Harm, 99
Brouwer, Susanne, 186
Burchfield, Laura, 251
Cai, Zhenguang, 121, 229
Cali, Maria Lucia, 225
Callaway, Frederick, 241
Cane, James, 58
Carlson, Katy, 25, 106, 108
Carreiras, Manuel, 187
Carroll, Rebecca, 104
Chambers, Craig, 54, 249
Chan, San, 174
Chang, Franklin, 31, 134, 136, 171
Chapman, Ashley, 81
Chemla, Emmanuel, 118
Chen, Baoguo, 60
Chernova, Daria, 124
Chiriacescu, Sofiana, 69, 259
Christianson, Kiel, 74
Christophe, Anne, 29, 35
Chu, Mingyan, 238
Chung, Eunyong, 243
Cimiano, Philipp, 37
Clifton, Charles, 82
Clostre, Nourane, 208
Coco, Moreno I., 163
Cohen Priva, Uriel, 176
Cokal, Derya, 71
Colombo, Lucia, 112
Colonna, Saveria, 68
Cooper, Angela, 80
Corley, Martin, 73, 188, 230
Costa, João, 226
Coulson, Seana, 244, 262
Covey, Lauren, 213
Crepaldi, Davide, 79, 113, 190, 202, 237
Cristante, Valentina, 137
Crocker, Matthew, 182, 185
<table>
<thead>
<tr>
<th>Author Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cummins, Chris</td>
<td>118, 119</td>
</tr>
<tr>
<td>Cunnings, Ian</td>
<td>160, 183</td>
</tr>
<tr>
<td>Damian, Markus</td>
<td>180, 250</td>
</tr>
<tr>
<td>Darby, Jeannique</td>
<td>157</td>
</tr>
<tr>
<td>Darley, Emily</td>
<td>250</td>
</tr>
<tr>
<td>Dautriche, Isabelle</td>
<td>29, 35</td>
</tr>
<tr>
<td>Davies, Harriet</td>
<td>148</td>
</tr>
<tr>
<td>de Carvalho, Alex</td>
<td>29</td>
</tr>
<tr>
<td>de Jong, Nivja H.</td>
<td>87</td>
</tr>
<tr>
<td>de la Fuente, Israel</td>
<td>254</td>
</tr>
<tr>
<td>De Letter, Miet</td>
<td>178</td>
</tr>
<tr>
<td>De Marco, Rocco</td>
<td>113</td>
</tr>
<tr>
<td>De Santis, Anais</td>
<td>113</td>
</tr>
<tr>
<td>Declerck, Mathieu</td>
<td>175</td>
</tr>
<tr>
<td>Degani, Tamar</td>
<td>88</td>
</tr>
<tr>
<td>Degno, Federica</td>
<td>237</td>
</tr>
<tr>
<td>Dell, Gary</td>
<td>61</td>
</tr>
<tr>
<td>Delogu, Francesca</td>
<td>182</td>
</tr>
<tr>
<td>Demberg, Vera</td>
<td>255</td>
</tr>
<tr>
<td>Deng, Ying</td>
<td>179</td>
</tr>
<tr>
<td>Devlin, Joseph T.</td>
<td>89</td>
</tr>
<tr>
<td>Dickins, Jonny</td>
<td>206</td>
</tr>
<tr>
<td>Dijkstra, Ton</td>
<td>199</td>
</tr>
<tr>
<td>Dillon, Brian W.</td>
<td>44</td>
</tr>
<tr>
<td>Dimitrova, Diana</td>
<td>238</td>
</tr>
<tr>
<td>Dobrovie-Sorin, Carmen</td>
<td>114</td>
</tr>
<tr>
<td>Drake, Eleanor</td>
<td>230</td>
</tr>
<tr>
<td>Drenhaus, Heiner</td>
<td>182, 185</td>
</tr>
<tr>
<td>Durlik, Joanna</td>
<td>256</td>
</tr>
<tr>
<td>Elsiepen, Emilia</td>
<td>114</td>
</tr>
<tr>
<td>Elman, Jeffrey L.</td>
<td>260</td>
</tr>
<tr>
<td>Engelmann, Felix</td>
<td>158</td>
</tr>
<tr>
<td>Evert, Stefan</td>
<td>76</td>
</tr>
<tr>
<td>Family, Neiloufar</td>
<td>247</td>
</tr>
<tr>
<td>Farmer, Thomas A.</td>
<td>245</td>
</tr>
<tr>
<td>Fedele, Emily</td>
<td>70</td>
</tr>
<tr>
<td>Fedorenko, Evelina</td>
<td>23, 129</td>
</tr>
<tr>
<td>Feher, Olga</td>
<td>20, 204</td>
</tr>
<tr>
<td>Feldman, Elana</td>
<td>241</td>
</tr>
<tr>
<td>Felser, Claudia</td>
<td>160</td>
</tr>
<tr>
<td>Ferguson, Heather</td>
<td>58, 136</td>
</tr>
<tr>
<td>Ferguson, Lauren</td>
<td>93</td>
</tr>
<tr>
<td>Fernandes, Bruno</td>
<td>226</td>
</tr>
<tr>
<td>Fernandes, Eunice</td>
<td>163</td>
</tr>
<tr>
<td>Ferreira, Fernanda</td>
<td>71</td>
</tr>
<tr>
<td>Ferreira, Victor</td>
<td>77, 135, 203</td>
</tr>
<tr>
<td>Filik, Ruth</td>
<td>84, 132, 148, 239, 240</td>
</tr>
<tr>
<td>Finlayson, Ian</td>
<td>230</td>
</tr>
<tr>
<td>Fiorentino, Robert</td>
<td>213</td>
</tr>
<tr>
<td>Fitz, Hartmut</td>
<td>31</td>
</tr>
<tr>
<td>Flecken, Monique</td>
<td>200, 201</td>
</tr>
<tr>
<td>Fodor, Janet Dean</td>
<td>107</td>
</tr>
<tr>
<td>Foltz, Anouschka</td>
<td>37</td>
</tr>
<tr>
<td>Foraker, Stephani</td>
<td>222</td>
</tr>
<tr>
<td>Franck, Julie</td>
<td>224, 225</td>
</tr>
<tr>
<td>Fraser, Adam</td>
<td>33</td>
</tr>
<tr>
<td>Frazier, Michael</td>
<td>75, 161, 162</td>
</tr>
<tr>
<td>Freunberger, Dominik</td>
<td>102, 192, 257</td>
</tr>
<tr>
<td>Frisson, Steven</td>
<td>55, 147</td>
</tr>
<tr>
<td>Fukumura, Kumiko</td>
<td>93, 103</td>
</tr>
<tr>
<td>Futrell, Richard</td>
<td>246</td>
</tr>
<tr>
<td>Gabriele, Alison</td>
<td>213</td>
</tr>
<tr>
<td>Gallée, Jeanne</td>
<td>23</td>
</tr>
<tr>
<td>Gambi, Chiara</td>
<td>130</td>
</tr>
<tr>
<td>Gaskell, Gareth</td>
<td>62</td>
</tr>
<tr>
<td>Gaspers, Judith</td>
<td>37</td>
</tr>
<tr>
<td>Gatt, Albert</td>
<td>128</td>
</tr>
<tr>
<td>Gelormini-Lezama, Carlos</td>
<td>68</td>
</tr>
<tr>
<td>Gerstenberg, Tobias</td>
<td>53</td>
</tr>
<tr>
<td>Gibson, Edward</td>
<td>34, 129, 246</td>
</tr>
<tr>
<td>Goldstein, Louis</td>
<td>248</td>
</tr>
<tr>
<td>Gollan, Tamar</td>
<td>181</td>
</tr>
<tr>
<td>Goodall, Grant</td>
<td>264</td>
</tr>
<tr>
<td>Gotzner, Nicole</td>
<td>120</td>
</tr>
<tr>
<td>Graff, Peter</td>
<td>34</td>
</tr>
<tr>
<td>Grama, Ileana</td>
<td>67</td>
</tr>
<tr>
<td>Grant, Margaret</td>
<td>44</td>
</tr>
<tr>
<td>Grillo, Nino</td>
<td>226</td>
</tr>
<tr>
<td>Groen, Margriet</td>
<td>127</td>
</tr>
<tr>
<td>Grove, Julian</td>
<td>227</td>
</tr>
<tr>
<td>Gruberg, Nick</td>
<td>135</td>
</tr>
<tr>
<td>Guasti, Maria Teresa</td>
<td>209</td>
</tr>
<tr>
<td>Guerra, Ernesto</td>
<td>45, 116</td>
</tr>
<tr>
<td>Hadley, Lauren</td>
<td>216</td>
</tr>
<tr>
<td>Hagoort, Peter</td>
<td>57, 94, 238</td>
</tr>
<tr>
<td>Haigh, Matthew</td>
<td>72</td>
</tr>
<tr>
<td>Häikiö, Tuomo</td>
<td>206</td>
</tr>
<tr>
<td>Hale, John</td>
<td>241</td>
</tr>
<tr>
<td>Hanink, Emily</td>
<td>196, 227</td>
</tr>
<tr>
<td>Harris, Jesse</td>
<td>25, 106</td>
</tr>
<tr>
<td>Hartman, Jeremy</td>
<td>34</td>
</tr>
<tr>
<td>Hartshorne, Joshua K.</td>
<td>53</td>
</tr>
<tr>
<td>Hartsuiker, Robert</td>
<td>42, 167, 178</td>
</tr>
<tr>
<td>Harvey, Nicole</td>
<td>148</td>
</tr>
<tr>
<td>Hasham, Aimen</td>
<td>194</td>
</tr>
<tr>
<td>Havron, Naomi</td>
<td>24</td>
</tr>
<tr>
<td>Heller, Daphna</td>
<td>54</td>
</tr>
<tr>
<td>Hemforth, Barbara</td>
<td>114, 254</td>
</tr>
<tr>
<td>Hendriks, Petra</td>
<td>56, 138</td>
</tr>
</tbody>
</table>
AUTHOR INDEX

Hilberink-Schulpen, Béryl, 199
Hintz, Florian, 85
Hirose, Yuki, 166, 179, 217, 218, 222
Hoeks, John, 99
Hofmeister, Philip, 231
Hong, Upyong, 223, 243
Hoshino, ichi, 142
Huettig, Falk, 45, 63, 85, 195
Husain, Samar, 169
Hwang, Heeju, 177
Hyöna, Jukka, 206

Ibañeze, Antonio, 187
Idrissi, Ali, 90
Isola, Phillip, 129
Ito, Aine, 73
Ivanova, Iva, 203

Jackson, Carrie, 139
Jacob, Gunnar, 154, 247
Jacobs, Cassandra L., 61
Jäger, Lena, 158
Janciauskas, Marius, 134
Janse, Esther, 155
Järvikivi, Juhani, 125
Jeffrey-Wilensky, Jaclyn, 241
Jeong, in, 223
Joanisse, Marc F., 245
Joanna, Moy, 146
Johnson, Maria, 194
Jones, John Matthew, 208
Jones, Manon, 98
Julinková, Radka, 66

Kaiser, Elsi, 52, 70, 235, 248
Kamide, Yuki, 150, 151, 189
Katsika, Kalliopi, 168, 247
Kazanina, Nina, 27, 197, 198, 250
Kehayia, Eva, 95
Kempe, Vera, 175
Kentner, Gerrit, 111
Kerkhoff, Annemarie, 67
Khan, Manizeh, 123, 263
Kholodova, Alina, 247
Kim, Albert, 51
Kim, Jungho, 210
Kinsella, Jamal, 72
Kirby, Simon, 59, 204
Kirkici, Bilal, 154
Kirk, Neil, 175
Kirschenbaum, Amit, 64, 65
Kleinman, Daniel, 77, 181

Kliegl, Reinhold, 165
Kluender, Robert, 242, 244
Knoeferle, Pla, 116, 149
Köder, Franziska, 144
Koizumi, Masatoshi, 210
Konopka, Agnieszka, 96, 170, 171
Kotzor, Sandra, 172
Kowalski, Alix, 164
Krahmer, Emiel, 128
Krajewski, Grzegorz, 146
Kras, Tihana, 47
Kreiner, Hamutal, 88
Kretzschmar, Franziska, 100, 191
Krž, Adam, 205
Kukona, Anuenue, 150, 151
Kulakova, Eugenia, 102
Kunkel, Angelika, 239
Künert, Aylin, 30
Kush, Dave, 21
Kutas, Marta, 242, 260

Lago, Sol, 152
Lahiri, Aditi, 172
Laka, Itziar, 103
Lapesa, Gabriella, 76
Lau Zhu, Alex, 208
Lau, Ellen, 48, 152
Lee, Dongsu, 223
Lee, Yoonjeong, 248
Lelonkiewicz, Jaroslaw, 130
Lennon, Robert, 156
Leuthold, Hartmut, 148, 239
Lev-Ari, Shiri, 117
Levy, Roger, 153
Lindsay, Shane, 189
Lundquist, Bjorn, 188
Lüthi, Janine, 36
Lutz, David, 241

Mackenzie, Ian G., 239
Mády, Katalin, 22
Mahar, Adam, 241
Mahowald, Kyle, 34, 35, 129, 246
Marelli, Marco, 113, 190, 202
Martin, Andrea E., 50, 183
Martin, Clara, 187
Mattys, Sven, 22
Mauner, Gail, 122, 222
McElree, Brian, 224
McQueen, James, 126, 127
Meng, Michael, 193

268
Menon, Mythili, 235
Messenger, Katherine, 211
Metusalem, Ross, 260
Metzner, Paul, 101
Meyer, Antje, 85, 126, 155
Michel, Dan, 244
Milin, Petar, 97
Miller, Sam, 230
Minagawa, Yasuyo, 142
Mirkovic, Jelena, 62
Mitterer, Holger, 38, 105
Moers, Cornelia, 155
Mohaghegh, Mercedeh, 249
Molnar, Monika, 187
Monaghan, Padraic, 63
Morgan, Adam, 233
Morone, Elena Angela, 113
Mous, Maaren, 265
Murray, Wayne, 46
Myachykov, Andriy, 81
Myslin, Mark, 153

Nair, Vasavan P., 95
Nakamura, Chie, 217, 218
Nam, Yunju, 243
Naoi, Nozomi, 142
Nella Carminati, Maria, 149
Nicenboim, Bruno, 165
Nichols, Emily S., 245
Nieuwland, Mante S., 50, 141, 183, 261

O'Donnell, Tim, 53
Oines, Leif, 51
Oliva, Aude, 129
Omaki, Akira, 26, 219, 220
Opitz, Andreas, 64, 78
Özge, Duygu, 30
Özkan Gökşürk, Deniz, 30
Özyürek, Asli, 238

Pagan, Ascen, 206
Pappas, Leah, 139
Pechmann, Thomas, 78
Peperkamp, Sharon, 40
Perniss, Pamela, 174
Philalithis, Eugene, 131
Philip, Andrea, 175
Phillips, Colin, 48, 152
Piantadosi, Steven, 35
Pickering, Martin, 73, 141, 216, 230
Pine, Julian, 136
Pitiot, Alain, 240

Poeppel, David, 43
Poesio, Massimo, 184
Pohl, Muna, 143
Politzer-Ahles, Stephen, 41, 91
Porretta, Vincent, 125
Potter, David, 161
Potton, Anita, 194

Qian, Jingjing, 159
Qu, Qingqing, 180
Quené, Hugo, 87
Quinn, Katie, 93

Radulescu, Silvia, 207
Ralph-Nearman, Christina, 132, 240
Ramchand, Gillian, 188
Rasilo, Heikki, 212
Rees, Alice, 147
Reinisch, Eva, 38, 105
Riach, Calum, 83
Rice, Amy, 174
Ritt, Nikolaus, 20
Rizzi, Luigi, 225
Roberts, Adam C., 172
Rodd, Jennifer, 89, 229
Roehm, Dietmar, 102, 257
Rogahn, Maria, 64, 65
Roland, Douglas, 166, 179, 222
Rommers, Joost, 195
Ros, Idoia, 103
Rosa, Elise, 253
Rösler, Frank, 101
Rossomondo, Amy, 213
Rowland, Caroline, 136
Ruigendijk, Esther, 104, 138
Runnqvist, Elin, 77

Sabourin, Laura, 133
Sanders, Ted, 87
Santesteban, Mikel, 103, 215
Santi, Andrea, 226
Santiago, Julio, 208
Scheepers, Christoph, 109, 115, 189
Schiller, Niels, 265
Schimke, Sarah, 68
Schlesewsky, Matthias, 100
Schlueter, Zoe, 152
Schlueter, Kevin, 41
Schmid, Monika, 186
Schmidtke, Jens, 173
Schoot, Lotte, 57
Schouwenaars, Atty, 138
<table>
<thead>
<tr>
<th>Author Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schroeder, Sascha</td>
<td>206, 252</td>
</tr>
<tr>
<td>Schrötter, Pauline</td>
<td>145</td>
</tr>
<tr>
<td>Schuerman, William</td>
<td>126</td>
</tr>
<tr>
<td>Schulte im Walde, Sabine</td>
<td>76</td>
</tr>
<tr>
<td>Scott-Brown, Ken</td>
<td>175</td>
</tr>
<tr>
<td>Segaert, Katrien</td>
<td>57, 94</td>
</tr>
<tr>
<td>Seton, Bregtje</td>
<td>186</td>
</tr>
<tr>
<td>Shaoul, Cyrus</td>
<td>97</td>
</tr>
<tr>
<td>Shattuck-Hufnagel, Stefanie</td>
<td>110</td>
</tr>
<tr>
<td>Sheng, Yanan</td>
<td>49, 86</td>
</tr>
<tr>
<td>Silvey, Catriona</td>
<td>59</td>
</tr>
<tr>
<td>Singh, Raj</td>
<td>118</td>
</tr>
<tr>
<td>Siloussar, Natalia</td>
<td>124</td>
</tr>
<tr>
<td>Sloggett, Shayne</td>
<td>44, 152</td>
</tr>
<tr>
<td>Smith, Alastair</td>
<td>63</td>
</tr>
<tr>
<td>Smith, Cybelle</td>
<td>48</td>
</tr>
<tr>
<td>Smith, Garrett</td>
<td>32</td>
</tr>
<tr>
<td>Smith, Kenny</td>
<td>20, 59, 204</td>
</tr>
<tr>
<td>Smolik, Filip</td>
<td>205</td>
</tr>
<tr>
<td>Snedeker, Jesse</td>
<td>28, 30, 123, 263</td>
</tr>
<tr>
<td>Sorace, Antonella</td>
<td>47, 140, 188</td>
</tr>
<tr>
<td>Spalek, Antonella</td>
<td>120, 139</td>
</tr>
<tr>
<td>Staub, Adrian</td>
<td>100, 191</td>
</tr>
<tr>
<td>Stenneken, Prisca</td>
<td>37</td>
</tr>
<tr>
<td>Stepanov, Arthur</td>
<td>83</td>
</tr>
<tr>
<td>Sternewald, Wolfgang</td>
<td>258</td>
</tr>
<tr>
<td>Stewart, Andrew</td>
<td>72</td>
</tr>
<tr>
<td>Stowe, Laurie</td>
<td>186</td>
</tr>
<tr>
<td>Stranahan, Laine</td>
<td>263</td>
</tr>
<tr>
<td>Sturt, Patrick</td>
<td>47, 71, 183, 216</td>
</tr>
<tr>
<td>Su, Chung-I Erica</td>
<td>122</td>
</tr>
<tr>
<td>Sugiura, Masatoshi</td>
<td>236</td>
</tr>
<tr>
<td>Sulpiţio, Simone</td>
<td>112</td>
</tr>
<tr>
<td>Sun, Yue</td>
<td>40</td>
</tr>
<tr>
<td>Szewczyk, Jakub</td>
<td>232, 256</td>
</tr>
<tr>
<td>Tabor, Whitney</td>
<td>32</td>
</tr>
<tr>
<td>Tan, Rebecca</td>
<td>89</td>
</tr>
<tr>
<td>Tao, Yuan</td>
<td>184</td>
</tr>
<tr>
<td>Tenenbaum, Joshua</td>
<td>53</td>
</tr>
<tr>
<td>Theakston, Anna</td>
<td>146</td>
</tr>
<tr>
<td>Thiele, Kristina</td>
<td>37</td>
</tr>
<tr>
<td>Thierry, Guillaume</td>
<td>98</td>
</tr>
<tr>
<td>Thompson, Dominic</td>
<td>148</td>
</tr>
<tr>
<td>Tiffin-Richards, Simon</td>
<td>252</td>
</tr>
<tr>
<td>Ting Huang, Yi</td>
<td>164</td>
</tr>
<tr>
<td>Tjiong, Jade</td>
<td>87</td>
</tr>
<tr>
<td>Todd Jones, Jenna</td>
<td>198</td>
</tr>
<tr>
<td>Tooley, Kristen</td>
<td>170</td>
</tr>
<tr>
<td>Torabi Asr, Fatemeh</td>
<td>255</td>
</tr>
<tr>
<td>Tschirner, Erwin</td>
<td>64, 65</td>
</tr>
<tr>
<td>Tsegaye, Mulugeta</td>
<td>265</td>
</tr>
<tr>
<td>Tucker, Benjamin V.</td>
<td>125</td>
</tr>
<tr>
<td>Tucker, Matthew</td>
<td>90, 91</td>
</tr>
<tr>
<td>Tung, Mai</td>
<td>188</td>
</tr>
<tr>
<td>Turcan, Alexandra</td>
<td>84, 148, 240</td>
</tr>
<tr>
<td>Turk, Alice</td>
<td>110</td>
</tr>
<tr>
<td>Ulrich, Rolf</td>
<td>258</td>
</tr>
<tr>
<td>Urbach, Thomas P.</td>
<td>260</td>
</tr>
<tr>
<td>Uzundağ, Berna</td>
<td>30</td>
</tr>
<tr>
<td>van Bergen, Geertje</td>
<td>199, 200</td>
</tr>
<tr>
<td>Van de Cavey, Joris</td>
<td>167</td>
</tr>
<tr>
<td>van Deemter, Kees</td>
<td>128</td>
</tr>
<tr>
<td>van Gompel, Roger</td>
<td>128</td>
</tr>
<tr>
<td>Van Hell, Janet</td>
<td>139</td>
</tr>
<tr>
<td>van Rij, Jacolien</td>
<td>97</td>
</tr>
<tr>
<td>Van Rijn, Hedderik</td>
<td>56</td>
</tr>
<tr>
<td>Vasishth, Shravan</td>
<td>101, 158, 165, 169, 231</td>
</tr>
<tr>
<td>Vaughan-Evans, Awel</td>
<td>98</td>
</tr>
<tr>
<td>Vega-Mendoza, Mariana</td>
<td>140, 141</td>
</tr>
<tr>
<td>Vegh, Genna</td>
<td>196</td>
</tr>
<tr>
<td>Vergilova, Yoana</td>
<td>185</td>
</tr>
<tr>
<td>Vernice, Mirta</td>
<td>209</td>
</tr>
<tr>
<td>Vigliocco, Gabriella</td>
<td>174, 208</td>
</tr>
<tr>
<td>Villata, Sandra</td>
<td>224</td>
</tr>
<tr>
<td>Vinerte, Santa</td>
<td>133</td>
</tr>
<tr>
<td>Vinson, David</td>
<td>174, 208</td>
</tr>
<tr>
<td>Vitello, Sylvia</td>
<td>89</td>
</tr>
<tr>
<td>Vitkovitch, Melanie</td>
<td>194</td>
</tr>
<tr>
<td>Vogelzang, Margreet</td>
<td>56</td>
</tr>
<tr>
<td>von der Malsburg, Titus, 101</td>
<td></td>
</tr>
<tr>
<td>von Heusinger, Klaus</td>
<td>69</td>
</tr>
<tr>
<td>Vorwerg, Constanze</td>
<td>36</td>
</tr>
<tr>
<td>Wagers, Matt</td>
<td>221</td>
</tr>
<tr>
<td>Wakeford, Laura</td>
<td>46</td>
</tr>
<tr>
<td>Walbert, Kelly</td>
<td>201</td>
</tr>
<tr>
<td>Wang, Jessica</td>
<td>55</td>
</tr>
<tr>
<td>Wang, Lin</td>
<td>238</td>
</tr>
<tr>
<td>Wardlow, Liane</td>
<td>135</td>
</tr>
<tr>
<td>Wartenburger, Isabell</td>
<td>120</td>
</tr>
<tr>
<td>Watson, Duane</td>
<td>170</td>
</tr>
<tr>
<td>Weatherholtz, Kodi</td>
<td>39</td>
</tr>
<tr>
<td>Webman Shafran, Ronit</td>
<td>107</td>
</tr>
<tr>
<td>Weiβ, Anna Fiona</td>
<td>191</td>
</tr>
<tr>
<td>Wetterlin, Allison</td>
<td>172</td>
</tr>
<tr>
<td>Whan Cho, Allison</td>
<td>32</td>
</tr>
<tr>
<td>White, Laurence</td>
<td>22</td>
</tr>
<tr>
<td>Wijnen, Frank</td>
<td>67, 207</td>
</tr>
<tr>
<td>Williams, Glenn</td>
<td>151</td>
</tr>
<tr>
<td>Wilson, Colin</td>
<td>219</td>
</tr>
</tbody>
</table>
Wittenberg, Eva, 123
Wodniecka, Zofia, 256
Wood, Jeffrey, 72
Woollams, Anna, 146
Wu, Fuyun, 49, 86, 159
Wu, Qingqing, 263
Xiang, Ming, 196, 227
Yao, Panpan, 60
Yatabe, Kiyomi, 142
Yee Chow, Wing, 48, 152
Yoshida, Masaya, 27, 75, 161, 162, 214, 228
Yui Leung, Chi, 236
Yun, Hongoak, 223, 243
Yusa, Mayuko, 210
Yusa, Noriaki, 210
Zahn, Daniela, 109
Zahner, Katharina, 143
Zawiszewski, Adam, 215
Zhou, Peiyun, 74
Zhou, Xiaolin, 49