Online processing of scalar implicatures in Chinese as revealed by event-related potentials

Steve Politzer-Ahles[1], Robert Fiorentino[1], Xiaoming Jiang[2], Xiaolin Zhou[2]

[1]*University of Kansas,* [2]*Peking University*

18th Cognitive Neuroscience Society Annual Meeting (2011), San Francisco

Correspondence to: sjpa@ku.edu

Introduction

Terms like "some" have both *logical* and *pragmatic* readings:

1) "Some of the students there are hardworking."
   a) Logical: "Out of the set of students there, there exists some subset of students that are hardworking."
   b) Pragmatic: "Not all of the students there are hardworking."

Computing the pragmatic reading (1b) requires processing the speaker's communicative intention (s/he chose not to say "all", thus must have meant "not all"). (Katsos & Cummins, 2010; Noveck & Sperber, 2007; Tavano, 2010)

- Is any ERP component associated with processing the pragmatic meaning of terms like "some"?
- How quickly is the pragmatic meaning realized?
  - immediately (*default approach*, e.g. Levinson, 2000)
  - at a delay, after the logical meaning (*context-driven approach* e.g. Sperber & Wilson, 1995)?

Previous ERP studies

Only two previous ERP studies on scalar implicature processing

Noveck & Posada (2003)

- ERP responses to underinformative (logically correct but pragmatically infelicitous) sentence-final words
  2a) *Patently true*: "Some people have brothers."
  2b) *Patently false*: "Some couches have *windows."
  2c) *Underinformative*: "Some turtles have %shells."
- Underinformative sentences elicited a reduced N400 ERP component relative to other conditions.

Nieuwland et al. (2010)

- Addressed some methodological concerns from the previous study
- Compared underinformativeness effect to effect of lexico-semantic relatedness
  3a) *Informative/unrelated*: "Some people have pets..."
  3b) *Underinformative/related*: "Some people have #lungs…"
- Increased N400 for underinformative sentences only in subjs. with high pragmatic ability; N400 for unrelated sentences in other subjs.

Limitations

- Violations became apparent only downstream of quantifier.
  - No way to test whether pragmatic meaning of "some" was computed immediately when the quantifier was encountered
- Studies relied on real-world knowledge, perhaps initiating memory search for exceptions (shell-less turtle, pit-less cherry?).
- For some participants, underinformativess-related ERP is obscured by overlapping lexico-semantic N400 (Nieuwland et al., 2010).

Current study

Picture-sentence verification design (Wu & Tan, 2009; Tavano, 2010)

- Pictures provide controlled context for stimuli.
- Sentences are identical across violating and non-violating conditions.
  - No differences in lexico-semantic relatedness
- Violation becomes apparent as soon as quantifier "some" is read.
- Possible to compare effects of underinformativeness versus "patent falsehood" without introducing semantic incongruity

Mandarin Chinese as language of study

- Previous investigations have focused on Indo-European languages.

Design and method

Materials

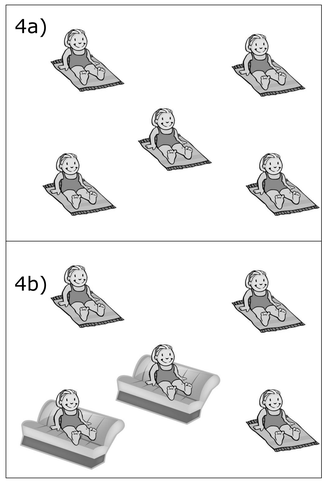
"All"-type (4a) and "some"-type (4b) pictures, matched with "all"-type (5a) and "some"-type (5b) sentences:

5a) *Túpiàn lǐ, suǒyǒu de nǚhái dōu zuòzài tánzi shang.*
"In the picture, all of the girls are sitting on blankets."

5b) *Túpiàn lǐ, yǒu de nǚhái zuòzài tánzi shang.*
"In the picture, some of the girls are sitting on blankets."

Conditions (2x2 design):

- (4a)+(5a): correct "all" sentence
- (4b)+(5a): logically incorrect "all" sentence"
- (4b)+(5b): correct "some" sentence
- (4a)+(5b): underinformative "some" sentence



Methods

- **Participants**: 9 native Chinese speakers from the University of Kansas
- **Stimuli:** 40 per condition, 148 fillers (74 correct "some" and "all", 37 incorrect object, 37 incorrect verb)
- **Procedure**: Picture displayed for 4000 ms, sentence displayed word-by-word (425 ms + 80 ms per character over 3; 400 ms ISI)
- **Task**: 10% of trials followed by judgment probe ("Is that correct?"), 10% by irrelevant comprehension question (e.g. "Are they wearing swimsuits?")
- **EEG recording**: EEG continuously recorded using Synamps2 amplifier (Compumedics Neuroscan, Inc.) and 32-channel Ag/AgCl electrode cap (Electro-Cap International, Inc.); data recorded at 1kHz with a 200Hz LPF/0.1Hz HPF
- **Data processing**: Artifacts (blinks, etc.) manually rejected; data re-referenced offline to average of both mastoids prior to epoching (-200 to 1000ms), baseline-correction, filtering, and averaging. ERPs were time-locked to the point at which the quantifier appeared.
- **Statistical analysis**: Calculated mean voltage amplitudes over selected time windows. Greenhouse-Geisser-corrected repeated measures ANOVAs with factors Type (pragmatic, logic), Violation (violation, no violation), scalp Anteriority (anterior, posterior), and scalp Laterality (left, midline, right) were performed.

Results

ERP results

- Relative to controls, both violation types elicited negativities from 150 to 900 msec.
- Based on waveforms and windowed ANOVAs, three time windows were identified:

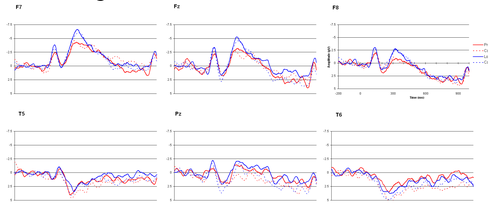


Figure 1) Grand average waveforms (15 Hz low-pass filter) at six sites. Negative is plotted upwards.

150 – 300 ms:

- Anteriority*Violation ($F(1,8)=11.627$, $p=0.009$): violating sentences elicited more negative waveforms in the posterior region

300 – 500 ms:

- Anteriority*Violation ($F(1,8)=12.196$, $p=0.008$): violating sentences more negative in posterior region (N400).
- Anteriority*Laterality*Type ($F(1.855, 14.836)=1.514$, $p=0.078$): "All" elicited negativity relative to "some"; effect broadly distributed anteriorly, left-central posteriorly

600-900 ms:

- Laterality*Type*Violation ($F(1.85, 14.802)=11.399$, $p=0.001$): Underinformative sentences elicited right-lateralized negativity, whereas logically incorrect sentences elicited broad negativity.
  - Effect of pragmatic violation greatest over right hemisphere
  - Effect of logic violation greatest over midline

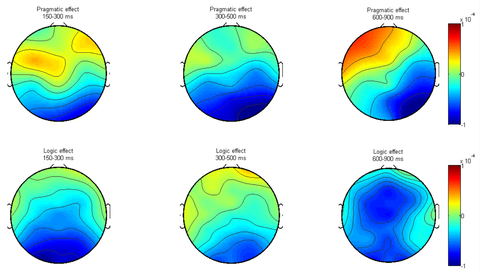


Figure 2) Scalp distributions of the pragmatic effect (underinformative – correct "some", top portion) and logic effect (logically false– correct "all", bottom portion) in three time windows

Results (cont'd)

Behavioral results

- 67% of underinformative sentences judged as correct (logical reading)
- 10% of logically incorrect sentences accepted; sig. less than underinformative acceptance rate, $t(7) = 4.252$, $p = .004$
- Greater proportion of logical readings than in previous studies (Tavano, 2010; Wu & Tan, 2009; De Neys & Schaeken, 2007; Bott & Noveck, 2004; Noveck & Posada, 2003)
- Likely due to presence of highly unacceptable fillers and to increased cognitive load (De Neys & Schaeken, 2007)
- 5 participants consistently made logical judgments; 1 consistently pragmatic; 2 inconsistent. (1 excluded because of response logging error)
  - In line with previous findings about consistency of responses to underinformativeness in experimental settings (Tavano & Kaiser, 2010)

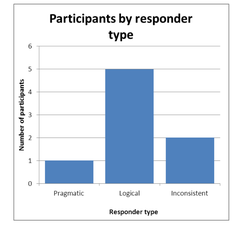


Figure 3) Number of participants by responder type

Discussion

Immediacy of implicature processing

- "Pragmatic reading" of *some* comes online immediately.
- Underinformativeness realizable on quantifier as well as on content word
- Consistent with findings from visual world eye-tracking (Tavano, 2010; Grodner et al., 2010; Degen & Tanenhaus, 2010; but see Huang & Snedeker, 2009, 2010)
- This finding is amenable to a default account of implicature generation, although it does not rule out a context-driven account.

Pragmatic versus logical processing

- Both processes elicit similar ERPs in the early (150-300 ms) and N400 (300-500 ms) time windows.
- Late negativity (600-900 ms) differentiates implicature-based and logical processing, showing a more right-lateralized effect for the former.
- Late effect seems to index more than just encountering an unexpected word (since both violating quantifiers should be unexpected).
- Late effect may index different processes initiated to cope with or make decisions about pragmatically and logically unlicensed quantifiers.

Understanding language-relevant brain responses

- N400-like effect can be elicited even on frequently repeated words, even though repetition reduces N400 component (Kutas & Federmeier, 2000).
- ERP responses to underinformative sentences can be elicited independently of lexico-semantic manipulations.
- Functional significance of the late negativity requires further research. It cannot only reflect conflict control in inference cancellation (see Pijnacker et al., 2011) since it was also observed in the logic (supposedly inference-free) condition.

Limitations

- Small sample; a larger replication of this study is currently under way
- High predictability of quantifier (no trials with other quantifiers; see Huang et al., 2010; Degen & Tanenhaus, 2010)
- The current design reveals electrophysiological activity associated with *violating* a scalar implicature, but not necessarily with *generating* one.

References

- Bonin, P; Peereman, R; Malardier, N; Méot, A; Chalard, M (2003). "A new set of 299 pictures for psycholinguistic studies:". *Behav. Res. Meth., Instr., & Comp.* 35(1), 158-67.
- Bott, L; Noveck, I (2004). "Some utterances are underinformative: the onset and time course of scalar inferences". *J. Mem. Lang.* 51, 437-57.
- Degen, J; Tanenhaus, K (2010). "When contrast is salient, pragmatic 'some' precedes logical 'some'". *23rd CUNY Conference on Human Sentence Processing.*
- De Neys, W; Schaeken, W (2007). "When people are more logical under cognitive load: Dual task impact on scalar implicature". *Exp. Psych.* 54(2), 128-33.
- Grodner, D; Klein, N; Carbary, K; Tanenhaus, M (2010). "Some, and possibly all, scalar inferences are not delayed: evidence for immediate pragmatic enrichment". *Cog.* 116, 42-55.
- Huang, Y; Hahn, N; Snedeker, J (2010). "Some inferences still take time: prosody, predictability, and the speed of scalar implicatures". *23rd CUNY Conf. on Human Sent. Proc.*
- Huang, Y; Snedeker, J (2009). "Online interpretation of scalar quantifiers: insight into the semantics-pragmatics interface". *Cog. Psych.* 58, 376-415.
- Katsos, N; Cummins, C (2010). "Pragmatics: from theory to experiment and back again". *Lang. & Ling. Compass* 4/5, 282-95.
- Kutas, M; Federmeier, K (2000). "Electrophysiology reveals semantic memory use in language comprehension". *Trends Cog. Sci.*4(12), 463-70.
- Levinson, S. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implications*. Cambridge, Massachusetts: MIT Press.
- Nieuwland, M; Ditman, T; Kuperberg, G (2010). "On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities". *J. Mem. Lang.* 63, 324-46.
- Noveck, I; Sperber, D (2007). "The why and how of experimental pragmatics: the case of scalar inferences". In *Advances in Pragmatics*.
- Noveck, I; Posada, A (2003). "Characterizing the time course of an implicature: an evoked potentials study". *Brain Res.* 85, 203-10.
- Pijnacker, J; Geurts, B; van Lambalgen, M; Buitelaar, J; Hagoort, P (2011). "Reasoning with exceptions: an event-related brain potentials study". *J. Cog. Nsci.* 23(2), 471-80.
- Sperber, D; Wilson, D (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Szekely, A; Jacobsen, T; D'Amico, S; Devescovi, A; et al. (2004). "A new on-line resource for psycholinguistic studies". *J. Mem. Lang.* 51(2), 247-50.
- Tavano, E (2010). *The Balance of Scalar Implicature.* Ph.D. thesis, University of Southern California.
- Tavano, E; Kaiser, E (2010). "Processing scalar implicature: what can individual differences tell us". *University of Pennsylvania Working Papers in Linguistics* 16(1).
- Urbach, T; Kutas, M (2010). "Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation". *J. Mem. Lang.* 63, 158-79.
- Wu, Z; Tan, J (2009). "Scalar implicature in Chinese child language: An experimental study". *Journal of Foreign Languages* 32(3), 69-75.

Acknowledgements: This research was funded by the National Science Foundation's East Asia and Pacific Summer Institutes (Award #1015160) and by the Center for Brain and Cognitive Sciences at Peking University. The pictures used in the study were adapted from Bonin et al. (2003), Szekely et al. (2004), or drawn by Peking University students.