Negotiating lexical uncertainty and expertise with disjunction

Roger Levy and Christopher Potts

Communicating in language about language

- Languages are neither fixed across time nor identically reproduced in all speakers, but rather continually renegotiated during interactions [7].
- People accommodate to each other's usage patterns [16], form temporarily lexical pacts [8, 3], and instruct each other about their linguistic views [18, 39].
- Some of this communication in language about language is direct, as with explicit definitions, but much of it arrives via secondary pragmatic inferences.
- Disjunction supports what appear to be opposing inferences about language:
  - **Hurfordian pressure [21]:** *X or Y* conveys that *X* and *Y* are disjoint
  - **Definitional inference [20]:** *X or Y* conveys that *X* and *Y* are synonymous
- This pattern is cross-linguistically robust, so we seek a single pragmatic model that can derive both of these meanings from the semantics of disjunction given different contextual assumptions.

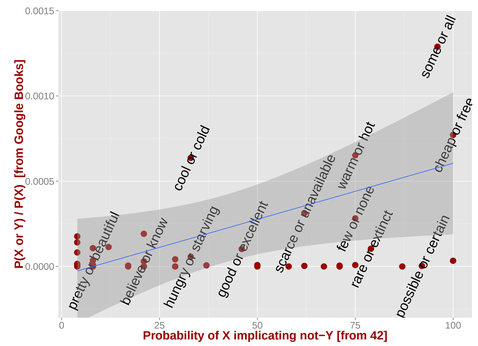
Hurfordian perceptions and intentions

Generalization: *X or Y* conveys that the speaker is using a lexicon in which *X* and *Y* are disjoint, or it addresses a speaker concern that the listener is using such a lexicon.

(1) the nuptials will take place in either France or Paris
(2) the canoe or boat will be held by the stream's current
(3) In 1940, 37% of us had gone to a church or synagogue in the last week.

No clear evidence for ordering restrictions or preferences deriving from the entailment relation:

Our corpus	
Disjunct order	Exs.
[general] or [specific]	79
[specific] or [general]	90



The frequency of *X or Y* correlates with the prevalence of *X* implicating *not Y* [5].

Disjunctive definition and identification

Generalization: *X or Y* can convey $[\![X]\!] \approx [\![Y]\!]$ when the speaker is mutually, publicly known to be an expert or would like to establish expertise.

(4) She's a wine lover or *oenophile*.
(5) Title: A Geological History of Manhattan or New York Island
(6) Welcome to New Haven or "the Elm City".
(7) It's a woodchuck, or land beaver.

- Motivation: speaker is a known 'instructor'; listener is a known non-expert.
- Motivation: speaker wishes to display expertise to another expert.
- Motivation: speaker sees value in (temporarily or permanently) defining a term.

Attested in Chinese, German, Hebrew, Ilokano, Japanese, Russian, and Tagalog. Seems to survive even where the language has a dedicated definitional disjunction morpheme (e.g., Finnish, Italian).

Further information

Paper, references, model code, corpus data: `http://github.com/cgpotts/pypragmods/`

Modeling communication with anxious experts

$$\cdots$$

$$S_2(m \mid w, \mathcal{L}) \propto \exp\left(\log\left(\alpha\, L_1(w \mid m, \mathcal{L})\right) - \beta \log\left(L_1(\mathcal{L} \mid m)\right) - C(m)\right)$$

world information − lexical preferences − costs

← Rich secondary meanings about language

$$L_1(w, \mathcal{L} \mid m) = l_1(w \mid m, \mathcal{L}) L_1(\mathcal{L} \mid m)$$

world information ∗ lexical discrimination

← suffices for manner implicature and embedded scalar implicature [38, 2]

$$l_1(w \mid m, \mathcal{L}) \propto s_1(m \mid w, \mathcal{L}) P(w)$$

← suffices for unembedded scalar implicature [12]

$$s_1(m \mid w, \mathcal{L}) \propto \exp\left(\log\left(\alpha\, l_0(w \mid m, \mathcal{L})\right) - C(m)\right)$$

← suffices for many kinds of ambiguity avoidance

$$l_0(w \mid m, \mathcal{L}) \propto \frac{\mathbb{I}(w \in \mathcal{L}(m))}{|\mathcal{L}(m)|} P(w)$$

← literal listener; interpretive semantics with priors

Definitional contexts

Require low disjunction costs and high β: the speaker is invested in communicating about the lexicon and can tolerate the cost of a disjunction that is synonymous with one of its disjuncts.

L_2 hears *A or X*	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_1, w_2\}]$	0	0	.08
$\mathcal{L}_1[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_2\}]$	.01	0	.08
$\mathcal{L}_2[A{:}\{\mathbf{w_1}\}, B{:}\{w_2\}, X{:}\{\mathbf{w_1}\}]$	.77	0	.06

$\alpha = 5$; $\beta = 7$; $C(or) = .01$

S_2 observes $\langle \mathcal{L}_2, w_1 \rangle$ ↓

A	*X*	*A or X*
.07	.48	.45

(bias against A *or* X is gone by S_3)

↓

L_1 hears *A or X*	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_1, w_2\}]$	0	0	.23
$\mathcal{L}_1[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_2\}]$	0	0	.38
$\mathcal{L}_2[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_1\}]$	.38	0	0

↙ ↓ ↘

l_1

$\mathcal{L}^*$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	.02	.02	.96
A or X	.02	.02	.96

$\mathcal{L}_1$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	0	1	0
A or X	.01	0	.99

$\mathcal{L}_2$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	1	0	0
A or X	1	0	0

↓ ↓ ↓

s_1

$\mathcal{L}^*$	*A*	*X*	*A or X*
w_1	.98	0	0
w_2	0	0	0
$w_1 \vee w_2$	0	.2	.2

$\mathcal{L}_1$	*A*	*X*	*A or X*
w_1	.99	0	0
w_2	0	.33	0
$w_1 \vee w_2$	0	0	.33

$\mathcal{L}_2$	*A*	*X*	*A or X*
w_1	.33	.33	.33
w_2	0	0	0
$w_1 \vee w_2$	0	0	0

↓ ↓ ↓

l_0

$\mathcal{L}^*$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	.33	.33	.33
A or X	.33	.33	.33

$\mathcal{L}_1$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	0	1	0
A or X	.33	.33	.33

$\mathcal{L}_2$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
X	1	0	0
A or X	1	0	0

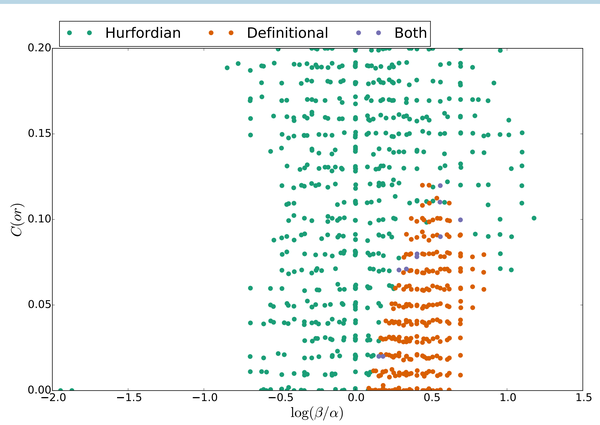
Hurfordian contexts

With high disjunction costs, exclusivization maximizes the justification for the long form; the Hurfordian instinct is a rational response to a disjunction that is unduly prolix for many lexica.

L_2 hears *A or X*	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_1, w_2\}]$	.02	0	.32
$\mathcal{L}_1[A{:}\{\mathbf{w_1}\}, B{:}\{w_2\}, X{:}\{\mathbf{w_2}\}]$	.04	0	.45
$\mathcal{L}_2[A{:}\{w_1\}, B{:}\{w_2\}, X{:}\{w_1\}]$	.03	0	.14

$\alpha = 2$; $\beta = 1$; $C(or) = 1$

Characterization



Summarizes a search over many parameter settings using a large lexicon and large world space.

References

[1] Leon Bergen, Noah D. Goodman, and Roger Levy. That's what she (could have) said: How alternative utterances affect language use. In *CogSci 34*, 2012.

[2] Leon Bergen, Roger Levy, and Goodman Noah D. Pragmatic reasoning through semantic inference. `http://web.mit.edu/bergen/www/papers/BergenLevyGoodman2014.pdf`, 2014.

[3] Susan E. Brennan and Herbert H. Clark. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1482–1493, 1996.

[4] Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, August 2004.

[5] Emmanuel Chemla. Apparent Hurford constraint obviations are based on scalar implicatures: An argument based on frequency counts. `http://www.emmanuel.chemla.free.fr/Material/Chemla-HurfordCounts.pdf`, 2013.

[6] Gennaro Chierchia, Danny Fox, and Benjamin Spector. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Maienborn et al. [30], pages 2297–2332.

[7] Herbert H. Clark. Dogmas of understanding. *Discourse Processes*, 23(3):567–59, 1997.

[8] Herbert H. Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22(1):1–39, 1986.

[9] Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st World Wide Web Conference*, pages 699–708, New York, 2012. ACM.

[10] Judith Degen and Michael Franke. Optimal reasoning about referential expressions. In *Proceedings of SemDIAL 2012*, Paris, September 2012.

[11] Brian Dillon, Charles Clifton, and Lyn Frazier. Pushed aside: Parentheticals, memory and processing. *Language, Cognition and Neuroscience*, 29(4):483–498, 2014.

[12] Michael C. Frank and Noah D. Goodman. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998, 2012.

[13] Michael Franke. *Signal to Act: Game Theory in Pragmatics*. ILLC Dissertation Series. Institute for Logic, Language and Computation, University of Amsterdam, 2009.

[14] Gerald Gazdar. *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press, New York, 1979.

[15] Howard Giles. Communication accommodation theory. In *Engaging Theories in Interpersonal Communication: Multiple Perspectives*, pages 161–173. Sage Publications, 2008.

[16] Howard Giles, Nikolas Coupland, and Justine Coupland. Accommodation theory: Communication, context, and consequences. In Howard Giles, Nikolas Coupland, and Justine Coupland, editors, *Contexts of Accommodation*. Cambridge University Press, 1991.

[17] Noah D. Goodman and Andreas Stuhlmüller. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5(1):173–184, 2013.

[18] Marti A. Hearst. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of COLING 1992*, pages 539–545, Nantes, August 1992. Association for Computational Linguistics.

[19] Julia Hirschberg. *A Theory of Scalar Implicature*. PhD thesis, University of Pennsylvania, 1985.

[20] Laurence R. Horn. *A Natural History of Negation*. Chicago, 1989.

[21] James R. Hurford. Exclusive or inclusive disjunction. *Foundations of Language*, 11(3):409–411, 1974.

[22] Gerhard Jäger. Game dynamics connects semantics and pragmatics. In Ahti-Veikko Pietarinen, editor, *Game Theory and Linguistic Meaning*, pages 89–102. Elsevier, Amsterdam, 2007.

[23] Gerhard Jäger. Game theory in semantics and pragmatics. In Maienborn et al. [30], pages 2487–2425.

[24] Justine T. Kao, Leon Bergen, and Noah D. Goodman. Formalizing the pragmatics of metaphor understanding. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, pages 719–724, Wheat Ridge, CO, July 2014. Cognitive Science Society.

[25] Justine T. Kao, Jean Y. Wu, Leon Bergen, and Noah D. Goodman. Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences*, August 2014.

[26] Stephen C. Levinson. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. MIT Press, Cambridge, MA, 2000.

[27] Roger Levy and Carl Pollard. Coordination and neutralization in HPSG. In Frank Van Eynde, Lars Hellan, and Dorothee Beermann, editors, *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, pages 221–234. CSLI Publications, 2001.

[28] David Lewis. *Convention*. Harvard University Press, Cambridge, MA, 1969.

[29] Mark Liberman. Questioning reality. `http://itre.cis.upenn.edu/~myl/languagelog/archives/001837.html`, January 2005.

[30] Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors. *Semantics: An International Handbook of Natural Language Meaning*, volume 3. Mouton de Gruyter, Berlin, 2012.

[31] Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014):176–182, 2011.

[32] Christopher Potts. *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford, 2005.

[33] Christopher Potts. Conventional implicature and expressive content. In Maienborn et al. [30], pages 2516–2536.

[34] Christopher Potts. Conversational implicature: Interacting with grammar. `http://web.stanford.edu/~cgpotts/manuscripts/potts-interacting2013.pdf`, 2013.

[35] Günther Rohdenburg. Dogs, bitches and other creatures. *Journal of Semantics*, 4(2):117–135, 1985.

[36] Thomas C. Schelling. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA, 1960.

[37] Raj Singh. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy*, 31(2):245–260, 2008.

[38] Nathaniel J. Smith, Noah D. Goodman, and Michael C. Frank. Learning and using language via recursive pragmatic reasoning about other agents. In *NIPS 26*, 2013.

[39] Rion Snow, Daniel Jurafsky, and Andrew Y. Ng. Learning syntactic patterns for automatic hypernym discovery. In Lawrence K. Saul, Yair Weiss, and Léon Bottou, editors, *Advances in Neural Information Processing Systems 17*, pages 1297–1304, Cambridge, MA, 2005. MIT Press.

[40] Alex Stiller, Noah D. Goodman, and Michael C. Frank. Ad-hoc scalar implicature in adults and children. In Laura Carlson, Christoph Hoelscher, and Thomas F. Shipley, editors, *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, pages 2134–2139, Austin, TX, July 2011. Cognitive Science Society.

[41] Kristen Syrett and Todor Koev. Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics*, 2014.

[42] Bob van Tiel, Emiel van Miltenburg, Natalia Zevakhina, and Bart Geurts. Scalar diversity. Ms., University of Nijmegen, 2013.

[43] Judith Tonhauser, David I. Beaver, Craige Roberts, and Mandy Simons. Towards a taxonomy of projective content. *Language*, 89(1):66–109, 2013.

[44] Adam Vogel, Max Bodoia, Christopher Potts, and Dan Jurafsky. Emergence of Gricean maxims from multi-agent decision theory. In *Human Language Technologies: The 2013 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 1072–1081, Stroudsburg, PA, June 2013. Association for Computational Linguistics.

[45] Adam Vogel, Andrés Gómez Emilsson, Michael C. Frank, Dan Jurafsky, and Christopher Potts. Learning to reason pragmatically with cognitive limitations. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, pages 3055–3060, Wheat Ridge, CO, July 2014. Cognitive Science Society.