The best response: Speaker rationality in an interactive paradigm

Anton Benz, Nicole Gotzner

Centre for General Linguistics (ZAS)
Berlin

Workshop on Rationality, Probability and Pragmatics
25.-27. May 2016

Section 1

Implicature in Complex Sentences

Scalar Implicature and Implicature of Complex Sentences

Example (Unembedded implicature trigger)

1. **Some** of the girls found marbles.
   $+>$ Not all found marbles.

Example (Embedded implicature trigger)

1. Every girl found **some** of their marbles.

The Standard Theory

Levinson (1983)

Example (Unembedded implicature trigger)

Some of the girls found marbles. $= A(\text{some})$
$+>$ Not all girls found marbles.

Reasoning:

- $A(\text{all})$: would have been more informative
- Speaker did't say $A(all)$ but $A(some)$
- Hence, he must believe $\neg A(some)$
- Cooperativity + competence $\Rightarrow \neg A(some)$

Horn scales: $\langle$*all, some*$\rangle$, $\langle$*and, or*$\rangle$,...

Embedded Implicature

Example

1. Every girl found some of her marbles.
2. Some of the girls found some of their marbles.

Variously predicted readings:

(Chierchia 2004, Sauerland 2004, Chierchia et al 20012)

1. Every girl found some of her marbles.
   — **literal**: Every girl found some and possibly all of her marbles.
   — **global**: Not every girl found all of her marbles.
   — **local**: Every girl found some but not all of her marbles.
2. Some of the girls found some of their marbles.
   — **glob/loc**: Not all of the girls found some of their marbles.
   — **glob**: None of the girls found all of her marbles.

Game Theory and Complex Sentences

Interactional approaches:

- Franke (2009) / Jäger (2013): Iterated Best Response Models.
- Benz (2012): Error Models.
- Pavan (2013) / Rothschild (2013): Iterated Admissibility Models.
- Bergen et al. (in print), Potts et al. (in print): Bayesian Models.

Problem:

- GT provides no technique for analysing linguistic structure.
- Seems to be confined to a globalist approach.

Aims of this Talk

- Present a specific model of implicature in complex sentences.
- Introduce new interactive experimental paradigm for testing the model.
- Evaluate experimental results/different speaker strategies.

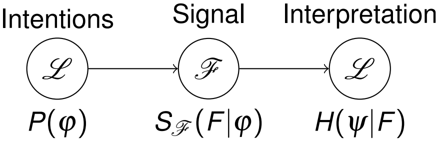
Section 2

Error Models

Error Models

Benz 2012

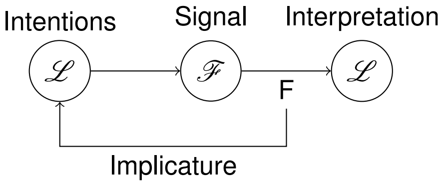
- Communication as stochastic process (Shannon 1948).



Error Models

Benz 2012

- Communication as stochastic process (Shannon 1948).



- If hearer can uniquely recover intended message ⇒ Success.
- If not ⇒ Clarification request.

Error Models: Critical Example

The Role of Errors

Example (Bus Ticket)

An email was sent to all employees that bus tickets for a joint excursion have been bought and are ready to be picked up. By mistake, no contact person was named. Hence, *H* asks one of the secretaries:

H: Where can I get the bus tickets for the excursion?

S: Ms. Müller is sitting in office 2.07. ($U_{M2.07}$)

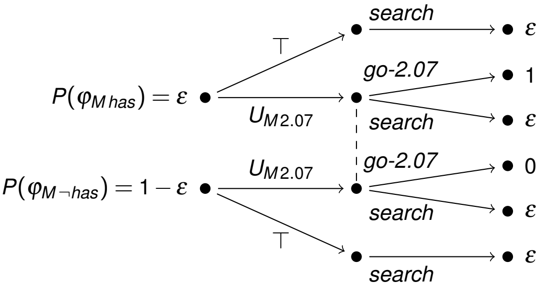
$+>$ Bus tickets are available from Ms. Müller.

Problem:

- Hearer finds list with all room numbers of all employees.
  $\not\Rightarrow$ Goes to 2.07.

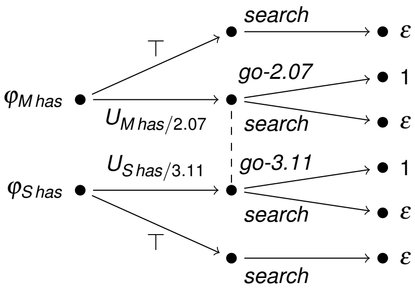
$\Rightarrow$ Literal content not enough for inducing hearer to choose intended action.

A Game Tree



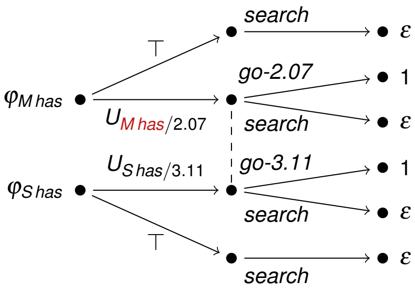
- **Problem**: $EU(go\text{-}2.07|[\![U_{M\,2.07}]\!]) = EU(search|[\![U_{M\,2.07}]\!]) = \varepsilon$.
- **Implies**: Literal content is irrelevant.

How it should have been played



- Literal content is deciding optimal action.
- In Example speaker omitted part of message.

Omitting Part of Message



- Literal content is deciding optimal action.
- In Example speaker omitted part of message.

Presentation in Table

- Consider for each state of affairs the optimal assertions.
- Consider all utterances which can result from omitting a conjunct of message.
  $\Rightarrow$ Noise ($\mathscr{N}_\varphi$)
- Consider the reduced utterances from which the original message can be reconstructed. ($\mathscr{U}_\varphi$)

φ	$Lit(\varphi)$	$\mathscr{N}_\varphi$	$\mathscr{U}_\varphi$
$\varphi_{M\,has/2.07}$	$U_{M\,has/2.07}$	$U_{M\,has/2.07}, F_{M\,has}, U_{M\,2.07}$	$U_{M\,2.07}$
$\varphi_{M\,has/3.11}$	$F_{M\,has/3.11}$	$F_{M\,has/3.11}, F_{M\,has}, F_{M\,3.11}$	$F_{M\,3.11}$
$\varphi_{S\,has/2.07}$	$F_{S\,has/3.11}$	$F_{S\,has/3.11}, F_{S\,has}, F_{S\,3.11}$	$F_{S\,3.11}$
$\varphi_{S\,has/3.11}$	$U_{S\,has/3.11}$	$U_{S\,has/3.11}, F_{S\,has}, F_{S\,3.11}$	$F_{S\,3.11}$

Characteristics

- Uniform explanation of relevance and quantity implicature
- No gambling: short utterances communicate message with certainty
- No blocking: preference for short utterances does not lead to risky utterances
- No hidden semantic operators

Section 3

Testing for Implicature by Decision Making

The basic best response paradigm
Nicole Gotzner & Anton Benz

Scenario

Background:

- 4 girls who each own a set of 4 special edition marbles (Degen & Goodman, 2014);
- marbles get lost during play and girls have to find them
- mother offers rewards to girls

Reward system:

- chocolate: girl finds all 4 of her marbles
- candy: girl finds fewer than 4 of her marbles
- gummy bears: girl finds none of her marbles (consolation prize)

Instructions

- Mother tells participants how many marbles each girl found
- Task: Participants are asked to buy sweets for the girls

Example

Sentence: **No girl found any of her marbles**

Chocolate	☐	YES	*X*	No
Candy	☐	YES	*X*	No
Gummy bear	*X*	YES	☐	No

The Experiment as Signalling Game

Playing the game:

1. Mother = speaker knows actual world
2. Mother chooses an utterance
3. Subject chooses an action: buying sweets
4. Game ends

- Game structure common knowledge
- Game of pure coordination: preferences aligned

Preferences:

- Every girl should get her appropriate sweet
- No superfluous sweets should be bought

Seven possible worlds

- $\exists|\nexists$: Some found none
- $\exists|\exists^!$: Some found some but not all
- $\exists|\forall$: Some found all

$\exists\|\nexists$	$\exists\|\exists^!$	$\exists\|\forall$	world
1	0	0	$v_1 = \square_1$
0	1	0	$v_2 = \square_2$
0	0	1	$v_3 = \blacksquare_3$
1	1	0	$v_4 = \square_4$
1	0	1	$v_5 = \square_5$
0	1	1	$v_6 = \square_6$
1	1	1	$v_7 = \square_7$

Seven possible actions

Actions

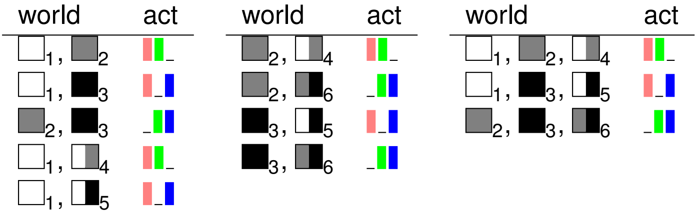
- ▮: Gummy bear
- ▮: Candy
- ▮: Chocolate

Best responses

world	act	world	act	world	act
$v_1 = \square_1$	▮ _ _	$v_4 = \square_4$	▮ ▮ _	$v_7 = \square_7$	▮ ▮ ▮
$v_2 = \blacksquare_2$	_ ▮ _	$v_5 = \square_5$	▮ _ ▮		
$v_3 = \blacksquare_3$	_ _ ▮	$v_6 = \blacksquare_6$	_ ▮ ▮		

Effect of Uncertainty

Best responses



- In all other cases: 

Error Models for Complex Sentences

Different alternatives:

- Alternative utterances constructed from worlds.

Literal descriptions of worlds:

world	utterances
1	$U_{\forall\|\nexists}$
	$U_{\nexists\|\exists}$
2	$U_{\forall\|\exists!}$
	$U_{\forall\|\exists} \wedge U_{\nexists\|\forall}$
3	$U_{\forall\|\forall}$

world	utterances
4	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\exists!} \wedge U_{\nexists\|\forall}$
5	$U_{\exists!\|\nexists} \wedge U_{\nexists\|\exists!} \wedge U_{\exists!\|\forall}$
6	$U_{\forall\|\exists} \wedge U_{\exists!\|\forall} \wedge U_{\exists!\|\exists!}$
	$U_{\exists!\|\forall} \wedge U_{\exists!\|\exists!} \wedge U_{\nexists\|\nexists}$
7	$U_{\exists!\|\forall} \wedge U_{\exists!\|\exists!} \wedge U_{\exists!\|\nexists}$

- $U_{Q|Q'}$: Q of the girls found Q' of the marbles.
- $\exists^!$: some but not all, $\nexists$: none

Error Models for Complex Sentences

- Shorter utterances constructed by elimination rules

Elimination rules:

1. $U_{(\exists!)} \rightarrow U_{(\exists)}$: reduction of '*some but not all*' to '*some*'
2. $U_{\nexists|\alpha} \wedge U_{\beta} \rightarrow U_{\beta}$: elimination of conjuncts with empty subjects

Restrictions:

- Rule $U_x \rightarrow U_y$ only applicable if $[\![U_x]\!] \subseteq [\![U_y]\!]$
- Requirement: unique recoverability of meaning (long story)

Application of Elimination rules

Literal descriptions of worlds:

world	utterances
$_1$	$U_{\forall\|\nexists} \wedge U_{\nexists\|\exists!} \wedge U_{\nexists\|\forall}$
$_2$	$U_{\nexists\|\nexists} \wedge U_{\forall\|\exists!} \wedge U_{\nexists\|\forall}$
$_3$	$U_{\nexists\|\nexists} \wedge U_{\nexists\|\exists!} \wedge U_{\forall\|\forall}$
$_4$	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\exists!} \wedge U_{\nexists\|\forall}$
$_5$	$U_{\exists!\|\nexists} \wedge U_{\nexists\|\exists!} \wedge U_{\exists!\|\forall}$
$_6$	$U_{\nexists\|\nexists} \wedge U_{\exists!\|\exists!} \wedge U_{\exists!\|\forall}$
$_7$	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\exists!} \wedge U_{\exists!\|\forall}$

Application of Elimination rules

Elimination of ‘none of the girls . . . :

-

world	utterances
1	$U_{\forall\|\nexists}$
2	$U_{\forall\|\exists!}$
3	$U_{\forall\|\forall}$
4	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\exists!}$

world	utterances
5	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\forall}$
6	$U_{\exists!\|\exists!} \wedge U_{\exists!\|\forall}$
7	$U_{\exists!\|\nexists} \wedge U_{\exists!\|\exists!} \wedge U_{\exists!\|\forall}$

Application of Elimination rules

Elimination of some but not all:

world	utterances
$\square_1$	$U_{\forall\|\nexists}$
$\square_2$	$U_{\forall\|\exists}$
$\square_3$	$U_{\forall\|\forall}$
$\square_4$	$U_{\exists\|\nexists} \wedge U_{\exists\|\exists}$

world	utterances
$\square_5$	$U_{\exists\|\nexists} \wedge U_{\exists\|\forall}$
$\square_6$	$U_{\exists\|\exists} \wedge U_{\exists\|\forall}$
$\square_7$	$U_{\exists\|\nexists} \wedge U_{\exists\|\exists} \wedge U_{\exists\|\forall}$

Application of Elimination rules

Predicted maximal simplification:

-

world	utterances
$\square_1$	$U_{\forall\mid\nexists}$
$\blacksquare_2$	$U_{\forall\mid\exists}$
$\blacksquare_3$	$U_{\forall\mid\forall}$
$\square_4$	$U_{\exists\mid\nexists} \wedge U_{\exists\mid\exists}$

world	utterances
$\square_5$	$U_{\exists\mid\nexists} \wedge U_{\exists\mid\forall}$
$\blacksquare_6$	$U_{\exists\mid\exists} \wedge U_{\exists\mid\forall}$
$\square_7$	$U_{\exists\mid\nexists} \wedge U_{\exists\mid\exists} \wedge U_{\exists\mid\forall}$

Testing the Model

Testable predictions:

- utterance length increases with complexity of world.
- critical strategy not less efficient than average human strategy.
- strategy is efficient:
  - — increasing average utterance length does not increase communicative success.
  - — decreasing average length should decrease communicative success (??) (claim in general probably not correct, however, we expected it in marble scenario)

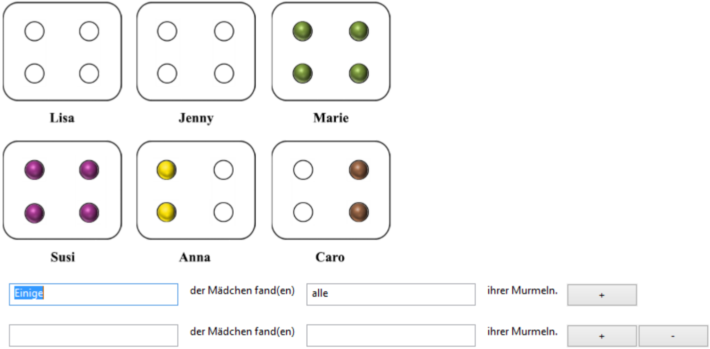
Section 4

The interactive best response paradigm

Anton Benz & Nicole Gotzner

The interactive best response paradigm

- Participants play best response paradigm in groups, taking two different roles (speaker and hearer)
- **Speaker's task**: Describe state of the world represented by picture
- **Response options**: *all, some, none, some but not all, some and possibly all, any* (up to 5 sentences)



Hearer's task

- **Hearer's Task**: Buy sweets based on speaker's description

  'Some of the girls found all of their marbles and some of the girls found none of their marbles.'

chocolate	☐	YES	☐	NO
candy	☐	YES	☐	NO
gummy bear	☐	YES	☐	NO

Methods

Procedure:

- Participants are randomly assigned to each role (3 times)
- System pairs two participants, pairings change across blocks
- Participants learn reward system with pictures in practice phase

Items:

- 7 worlds are instantiated by six items
- In one block, a world is shown only once

Participants:

- 38 German participants (mean age: 29.3, 21 female, 17 male)
- 2 groups with 4 players (8), 5 groups with 2 players (10)
- 6 groups with 3 players plus experimenter (18); experimenter employs critical strategy (produces statements predicted by Benz' model)

Success rate of utterance

Calculation of success rate:

- Use average response/interpretation of participants
- data of experimenter are eliminated
- $p_i(w|u)$: probability of participant i interpreting utterance u as w

$$Succ(u|w) = \sum_{i \in H(u)} p_i(w|u)/|H(u)|,$$

$H(u)$: set of participants who interpreted utterance u.

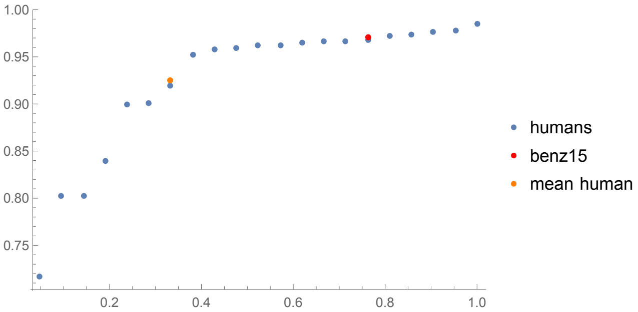
Results: Success rate and length of utterance

Results (→: critical strategy):

	utterances	world	success %	length
→	(None–any)	1	98%	1.0
	(All–none)		100%	
→	(All–some)	2	94%	1.58
	(All–some but not all)		93%	
→	(All–all)	3	99%	1.0
→	(Some–some, some–none)	4	95%	2.72
	(Some–none, some–some, none–all)		100%	
	(Some–some)		25%	
→	(Some–all, some–none)	5	96%	2.63
→	(Some–all, some–some)	6	98%	2.60
→	(Some–all, some–some, some–none)	7	100%	3.27

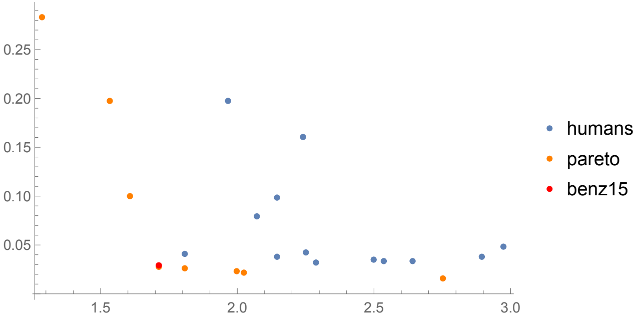
Results: Comparison with individual strategies

- Success rates of individual players with utterances occurring more than once in corpus
- Critical (Benz' model): 0.971, average participant: 0.925
- Critical strategy is significantly better than average participant strategy (one-tailed t test: $p < .001$)



Results: Comparison with individual strategies

- y-axis: 1− Success rates of individual players (utterances occurring more than once).
- x-axis: average length of utterances of strategy
- Critical: av.–length: 1.71429, failure rate: 0.029



Section 5

Comparison with other Theories

Structural Accounts

Sentence level accounts:

	Chierchia 2004	Sauerland 2004	Observed
some some	$\{\square_4, \square_7\}$	$\{\square_4\}$	$\square_4$ (25%), $\square_2$ (75%)
some some & some all	$\{\square_7\}$	$\{\square_5, \square_7\}$	$\square_6$ (98%)
some none & some some & some all	$\{\square_7\}$	$\{\square_5, \square_7\}$	$\square_7$ (100%)
some none & some all	$\{\square_5, \square_7\}$	$\{\square_5, \square_7\}$	$\square_5$ (96%)

Predictions of Modern Localism

Example:

- Conjunction: Some some and some all
- Observed interpretation: $_6$ (98%)

Example (Modern Localism: Chierchia et al 2012)

possible readings

some	some	and some all	$_3$, $_4$, $_6$, $_7$
some	O [some]	...	$_6$, $_7$
O [some]	some	...	$_5$, $_7$
O [some]	O [some]	...	$_6$, $_7$
O [some	some]	...	$\perp$
O [some	some	...]	$_5$, $_7$

A Bayesian Model

Take into account:

- Aims at explanatory models (why vs. what)
- Uncertainty about contextual parameters
- Error prone communication
- Fit parametric models to data

What we tested:

- Pick one specific model: (Qing & Franke 2014)
- Fitted to experimental data
- Interested in qualitative behaviour

Reference Game Task

Speaker:

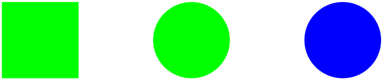
- chooses object: e.g. *green circle*
- signals: square, circle, green, blue

Hearer:

- receives signal, e.g. *green*
- chooses object, e.g. *green circle*

goal: speaker and hearer choose same object

An experimental token



Frank Goodman (2012), Qing & Franke (2015), Franke & Jäger (to appear)

Possible Parameters Influencing Production

- Probability with which hearer chooses worlds
- Preference for short utterances

Towards a model: Start with naive interpreter

Probability of choosing world w given utterance u:

$$P_{literal}(w|u) = \begin{cases} \frac{1}{|u|} & \text{if } w \in u \\ 0 & \text{else} \end{cases}$$

Expected utility of utterance u given w (disregarding preferences for signals):

$$EU(u|w) = P_{literal}(w|u).$$

Expected utility of utterance u given w (including preferences for signals):

$$EU(u|w) = P_{literal}(w|u) + cost(u). \quad (\text{if } w \in u)$$

Determining Speaker Production Probability

Background: Discrete rational choice theory.

- penalty *cost* for choosing colour: $0 \geq cost \geq -1$.
- degree of rationality λ

$$P_{prod}(u|w, \lambda, cost) = \frac{\exp(\lambda \cdot EU(u|w, cost))}{\sum_{u'} \exp(\lambda \cdot EU(u'|w, cost))}$$
$$= \frac{\exp(\lambda \cdot (P_{literal}(w|u) + cost(u)))}{\sum_{u'} \exp(\lambda \cdot (P_{literal}(w|u') + cost(u')))}$$

$$P_{inter}(w|u; \lambda, cost) = \frac{P(w)\, P_{prod}(u|w, \lambda, cost)}{\sum_{w'} P(w')\, P_{prod}(u|w', \lambda, cost)}$$

Two Models

Model 0:

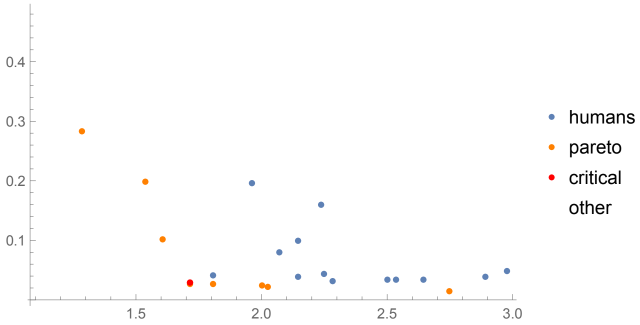
- $\lambda = 4.96$, $cost = -0.27$ (Pearson's ρ: 0.82)
- Fitted to average human production strategy (utterances occurring more than once)
- Literal interpretation strategy
- Fitted: Euclidean distance
- av. length: 1.83201, success rate: 0.78

Model 1:

- $\lambda = 4.45$, $cost = -0.94$ (Pearson's ρ: 0.68)
- Fitted to average human interpretation strategy (utterances occurring more than once)
- λ and *cost* represent production strategy against literal interpretation strategy
- Fitted: Euclidean distance
- av. length: 1.10759, success rate: 0.52806

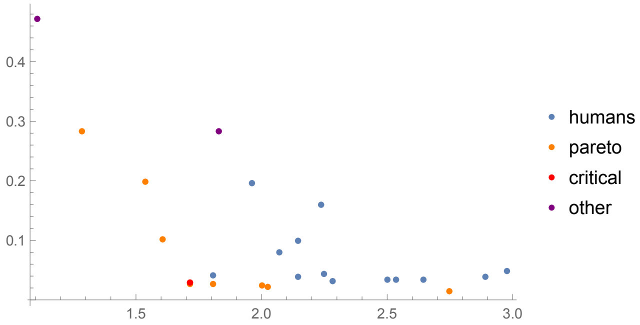
Comparison with individual strategies

- y-axis: 1− Success rates of individual players (utterances occurring more than once).
- X-axis: average length of utterances of strategy
- Critical: av.–length: 1.71429, failure rate: 0.029



Comparison with individual strategies

- y-axis: 1− Success rates of individual players (utterances occurring more than once).
- X-axis: average length of utterances of strategy
- Critical: av.–length: 1.71429, failure rate: 0.029



Conclusion

Results:

- utterance length increases with complexity of world.
- critical strategy more efficient than average human strategy.
- strategies with higher average utterance lengths not more successful.
- strategies with lower average utterance lengths less successful.
- Results pose problems for structural accounts (localism and globalism)

Future direction:

- Extend paradigm to more sentence types (downward entailing, disjunction, non–monotonic, and more)
- Look at relation to RSA-models.
- Study scenarios with partial speaker knowledge.

Thank you for your attention!

References I

[1] Anton Benz.
Errors in pragmatics.
Journal of Logic, Language, and Information, 21:97–116, 2012.

[2] Anton Benz.
Implicatures of complex sentences in error models.
In Andrea Schalley, editor, *Practical theories and empirical practice*, pages 273–306. John Benjamins, Amsterdam, 2012.

[3] Leon Bergen, Roger Levy, and Noah D. Goodman.
Pragmatic reasoning through semantic inference.
ms., 2014.

References II

[4] Gennaro Chierchia.
Scalar implicatures, polarity phenomena, and the syntax / pragmatics interface.
In Adriana Belletti, editor, *Structures and Beyond*, pages 39–103. Oxford University Press, Oxford, 2004.

[5] Gennaro Chierchia, Danny Fox, and Benjamin Spector.
Scalar implicature as a grammatical phenomenon.
In Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 3, pages 2297–2331. De Gruyter Mouton, Berlin, 2012.

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

References III

[7] Michael Franke.
Signal to Act: Game Theory in Pragmatics.
PhD thesis, Universiteit van Amsterdam, 2009.
ILLC Dissertation Series DS-2009-11.

[8] Michael Franke and Gerhard Jäger.
Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics.
To appear in *Zeitschrift für Sprachwissenschaft*, 2014.

[9] Herbert Paul Grice.
Logic and conversation.
In Peter Cole and Jerry L. Morgan, editors, *Syntax and Semantics*, volume 3, pages 41–58. Academic Press, New York, 1975.

References IV

[10] Herbert Paul Grice.
Studies in the Way of Words.
Harvard University Press, Cambridge MA, 1989.

[11] Gerhard Jäger.
Rationalizable signaling.
Erkenntnis, pages 1–34, 2013.

[12] Stephen C. Levinson.
Pragmatics.
Cambridge University Press, Cambridge, 1983.

[13] Sascia Pavan.
Quantity implicatures and iterated admissibility.
Linguistics and Philosophy, 36:261–290, 2013.

References V

[14] Christopher Potts, Daniel Lassiter, Roger Levy, and Michael C. Frank.
Embedded implicatures as pragmatic inferences under compositional lexical uncertainty.
Ms., Stanford University, 2015.

[15] Ciyang Qing and Michael Franke.
Variations on a bayesian theme: Comparing bayesian models of referential reasoning.
In Henk Zeevat and Hans-Christian Schmitz, editors, *Bayesian Natural Language Semantics and Pragmatics*, pages 201–220. Springer, Heidelberg, 2015.

[16] Daniel Rothschild.
Game theory and scalar implicatures.
Ms., All Souls College, Oxford, 2011.

References VI

[17] Uli Sauerland.
Scalar implicatures in complex sentences.
Linguistics and Philosophy, 27:367–391, 2004.

[18] Claude E. Shannon.
A mathematical theory of communication.
Bell System Technical Journal, 27:379–423, 623–656, 1948.