Natural language semantics and the psychology of reasoning

Salvador Mascarenhas

Ecole Normale Supérieure
Department of Cognitive Studies
Institut Jean-Nicod

Linguistic Analyses and Multilingualism
USiena, Spring 2019
Lecture #1

Language and reasoning

Take-home message

- Linguistic analysis **must** inform the study of how humans reason
- What we know about how we reason **must** inform our linguistic theories

Road map

- Lecture #1 (today)
  - A project for semantics as cognitive science
  - Reasoning and scalar implicature
- Research seminar (Thursday 2pm)
  - Human reasoners as hypothesis testers: don't just make true noises, tell stories
- Lecture #2 (Friday 9am/10am)
  - Reasoning in L2, fewer mistakes?
  - Reasoning with conditionals

Semantics and cognitive science

What are meanings?

When addressing the question "what is the meaning of X," we ask ourselves "what does it take for me to be able to say truthfully that I understand the meaning of X?"

Proper names

To know the meaning of "Donald J. Trump" I have to know that it refers to Donald J. Trump

Definite descriptions

To know the meaning of "the president of the United States" I have to know that it "refers" to whoever the president of the United States happens to be

Sentence meanings

The president of the United States is Hillary R. Clinton.

Observation

Speakers can classify an unbounded number of pairs of the form ⟨well-formed sentence, situation⟩ as yielding truth or falsehood

Truth-conditional semantics

To know the meaning of a sentence is to know under what conditions (i.e. situations) it is true (and under what conditions it is false)

Wait, is that it?

If two sentences are true in the same conditions, then by the definition above their meanings are identical

OK, sure

Not so sure...

Wait, is that it?

If two sentences are true in the same conditions, then by the definition above their meanings are identical

OK, sure

(1) Kim is a bachelor.

(2) Kim is an unmarried man.

Not so sure...

Wait, is that it?

If two sentences are true in the same conditions, then by the definition above their meanings are identical

OK, sure

(1) Kim is a bachelor.

(2) Kim is an unmarried man.

Not so sure...

(3) Two plus two equals four.
(4) Forty-five times six equals two-hundred and seventy.

(5) I want to take Ling Analyses & Multiling at USiena.
(6) I don't not want to take Ling Analyses & Multiling at USiena.

Wait, is that it?

If two sentences are true in the same conditions, then by the definition above their meanings are identical

OK, sure

(1) Kim is a bachelor.

(2) Kim is an unmarried man.

Not so sure...

(3) Two plus two equals four.
(4) Forty-five times six equals two-hundred and seventy.

(5) I want to take Ling Analyses & Multiling at USiena.
(6) I don't not want to take Ling Analyses & Multiling at USiena.

One way to put it

It seems like someone whose mind represents (3) wouldn't necessarily assent to (4).

Meaning, truth, and twins

So truth and truth conditions are perhaps not the whole story of what meanings are. Nevertheless, truth matters.

Twin earth

René was born in 1596 in France and died in 1650, on planet Earth. Imagine there were a planet Twin Earth that is exactly like earth in every respect, including its history, up to the year 1700 CE, except that water, instead of being dihydrogen monoxide, has some other chemical composition X. René and Twin René, being in their respective 17th centuries, have identical mental representations for the term "eau": dihydrogen monoxide and X are indistinguishable as far as they can tell. Yet René's "eau" and Twin René's twin "eau" seem to mean different things. (argument from Putnam, 1973)

It looks like there is some sense of "meaning" that is independent of what's going on in people's minds. That sense might well be exhausted by truth conditions

Meanings in the mind

- But clearly there's some other sense of "meaning" that depends on minds.
- "I understand the meaning of sentence X" suggests that there is something in my mind that corresponds to the meaning of sentence X
- In most contemporary traditions in cognitive science, we say that humans generate mental representations of the meanings of sentences

Semantics as a theory of what mental representations look like and how we arrive at them when interpreting linguistic utterances

Reasoning

A bat and a ball cost $1.10 together.
The bat costs $1.00 more than the ball.
How much does the ball cost?

Reasoning

A bat and a ball cost $1.10 together.
The bat costs $1.00 more than the ball.
How much does the ball cost?

- Human reasoning fails in strikingly systematic ways — we do something when we try to answer questions like these, but it looks like what we do is the wrong thing.
- Human reasoning works on mental representations that are often the product of interpretive processes (or always, in the case of reasoning experiments!) — linguistic processes feed into reasoning processes
- We will look at a number of interactions between interpretive processes and reasoning. We will study theories of the operations of the faculty for reasoning.

A project for semantics as cognitive science

- To focus on the overlap between semantics and the psychology of reasoning

Reasoning

- Reasoning is about naive entailment — whatever people deem to follow from a sentence of set of sentences
- Reasoning is about manipulating mental representations of sentences (among other things) to find new mental representations

- Semantics as cognitive science ought to produce representations of content appropriate to feed into a reasoning module and concern itself with what humans take to follow from sentences

Two notions of *what follows*

John speaks German and Mary speaks French, or else Bill speaks Italian.

Classical entailment (disjunctive syllogism)

If Bill doesn't speak Italian, then John speaks German.

Naive "entailment" (20/20 acceptance, Koralus & Mascarenhas, 2013)

If John speaks German, then Mary speaks French.

Two notions of representation of content

John speaks German and Mary speaks French, or else Bill speaks Italian.

Truth conditions

$(j \wedge m) \vee b$

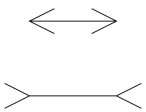
[worlds where John speaks German intersect worlds where Mary speaks French] union worlds where Bill speaks Italian

Inquisitive semantics (Mascarenhas, 2009, Groenendijk, 2008)

$\{j \wedge m, b\}$

two alternatives 1. John speaks German and Mary speaks French, 2. Bill speaks Italian

An analogy with vision



John speaks German
and Mary speaks French,
or Bill speaks Italian

Representation type			Desiderata for rep.	
	vision	**semantics**	**vision**	**semantics**
External	what is true of a piece of visual data	which worlds make a sentence true	horizontal lines have $=$ lengths	disjunctions are sets of worlds
Internal	what the human visual system constructs	what the human faculty of language constructs	horizontal lines have $\neq$ lengths	disjunctions raise alternatives

The psychology of reasoning

Failures of Reasoning

Lily pads

In a lake, there is a patch of lily pads.
Every day, the patch doubles in size.
It takes 48 days for the patch to cover the entire lake.
How long would it take for the patch to cover half of the lake?

Failures of Reasoning 2

Illusory inferences from disjunction

Mary has met every king or every queen of Europe.
Mary has met the king of the Netherlands.
Does it follow that Mary has met the king of Spain?

Failures of Reasoning 3

Married people

Jack is looking at Anne, but Anne is looking at George.
Jack is married, and George is not.
Is a married person looking at an unmarried person?

Yes No Cannot tell

Failures of Reasoning 3

Married people

Jack is looking at Anne, but Anne is looking at George.
Jack is married, and George is not.
Is a married person looking at an unmarried person?

Yes No Cannot tell

Anne is either married or she isn't.

Failures of Reasoning 4

Conjunction fallacy

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?

1. Linda is a bank teller.
2. Linda is a bank teller and is active in the feminist movement.

Conceptual housecleaning

Compelling fallacies

are (classically) invalid inference patterns that we often accept.

Repugnant validities

are (classically) valid inference patterns that we often reject.

Failures of *reasoning*?

- It's possible that some (perhaps many) of the problems we've seen so far are in fact the result of sound reasoning acting on non-obvious but perfectly reasonable and predictable interpretations of the premises
- If we don't countenance this possibility in a systematic and sophisticated way, we run the risk of misdiagnosing interesting interpretive processes as failures of reasoning

Affirming the consequent

P_1: If the card is long then the number is even.
P_2: The number is even.
Conclusion: The card is long.

Affirming the consequent

P_1: If the card is long then the number is even.
P_2: The number is even.
Conclusion: The card is long.

+ pragmatics

P_1: If the card is long then the number is even.
P_1^+: Only if the card is long is the number even.
P_2: The number is even.
Conclusion: The card is long.

Disjunction introduction

P_1: The card is long.
Conclusion: The card is long or the number is even.

Disjunction introduction

P_1: The card is long.
Conclusion: The card is long or the number is even.

+ pragmatics

$\varphi \vee \psi$ suggests strongly that one doesn't know which one of φ and ψ is in fact the case. But P_1 says that φ is the case! Compare with Moore's paradox:

It's raining but I don't believe it.
If it's raining and I don't believe it, then there is something I don't know.

Mental logic & mental models

- Mental logic and Mental models share the idea that humans intuitively do something related to formal logic when reasoning.
- Mental logic: humans do some form of proof theory, they have deductive rules they use to transform sentential representations into other sentential representations. E.g.: humans have a modus ponens rule *and* something like an affirming the consequent rule.
- Mental models: humans do some form of model theory, they build mental models (little "pictures") of the information they are given, update these representations sequentially, and inspect the resulting model for constitutive parts that are of interest.

Braine (1998), Johnson-Laird (1983)

Heuristics and biases

- We do not reason with logic or mathematics, instead we use heuristics that give quick and decisive answers to our questions. Our reasoning is rendered fallible by the imperfection of these heuristics, as well as a collection of innate biases.
- Judgment by substitution: Faced with a hard question, we substitute for it an easier question whose answers we can map into answers to the hard question. We work on answering the easier question (possibly with excellent accuracy!), and transpose the answer we find there to an answer to the harder question.
- Base-rate neglect: We often ignore general information about a representative sample and go purely on information about a particular individual.
- Anchoring effect: We will incorporate information we recognize as altogether irrelevant into our decision making.

Tversky & Kahneman (1974)

Probabilistic approach

- The functional aim of human reasoning is not to achieve logical validity. Reasoning is not simply an imperfect, quick & dirty way of approximating logic.
- Reasoning is about decision-making under uncertainty, it is intrinsically probabilistic. It proceeds by calculating posterior probabilities over possible conclusions, incorporating evidence using Bayes' rule.
- Logic is very conservative in handling uncertainty: if you have any doubt whatever, say no.
- What look like mistakes from a logical validity perspective are often in fact the most rational choices to make in the face of uncertainty in terms of Bayesian belief update.

Oaksford & Chater (2007)

Argumentative theory

- One question no one has tackled to anyone's satisfaction is that of *why and how* human reasoning evolved.
- Argumentative theory of reasoning: Reasoning evolved not to help us make better decisions, but to help us justify our decisions and persuade our interlocutors to adopt our views.
- Argumentation is an essential tool allowing for coordination of complex social activities, political structures, etc.
- Reasoning helps ground argumentation (loosely, but anything is better than nothing...).

Mercier & Sperber (2017)

Illusory inferences from disjunction

1. John speaks English and Mary speaks French, or else Bill speaks German.
   John speaks English.
   Mary speaks French?

Illusory inferences from disjunction(-like elements)

1. John speaks English and Mary speaks French, or else Bill speaks German.
   John speaks English.
   Mary speaks French?

2. Mary has met every king or every queen of Europe.
   Mary has met the king of the Netherlands.
   Mary has met the king of Spain?

3. Some pilot writes poems.
   John is pilot.
   John writes poems?

4. The gun fired and the guitar was out of tune, or else someone was in the attic.
   The trigger was pulled.
   The guitar was out of tune?

5. Each person at this party is either French or a linguist.
   That guy is European.
   That guy is French?

6. This party is only for French people and linguists.
   That guy is European.
   That guy is French?

7. John might speak English and French.
   John speaks English.
   John speaks French?

Illusory inferences from disjunction

P_1: Either Jane is kneeling by the fire and she is looking at the TV or otherwise Mark is standing at the window and he is peering into the garden.
P_2: Jane is kneeling by the fire.
Concl.: Jane is looking at the TV.

P_1: $(a \wedge b) \vee (c \wedge d)$
P_2: a
Concl.: b

- discovered by Walsh & Johnson-Laird (2004)
- 85%–90% acceptance, replicated multiple times
- Independent of how disjunction is expressed in English
- Original examples can be simplified significantly and made more natural

The erotetic theory of reasoning

Koralus & Mascarenhas (2013)

The erotetic principle

- *Part I* — Our natural capacity for reasoning proceeds by treating successive premises as questions and maximally strong answers to them. (problem of failure)
- *Part II* — Systematically asking a certain type of question as we interpret each new premise allows us to reason in a classically valid way. (problem of success)

Commitment on interpretation

Disjunctions raise alternatives and put pressure toward choosing an alternative — disjunctions are like questions in this regard (Inquisitive Semantics: Groenendijk, 2008, Mascarenhas, 2009)

Disjunctions are like questions

A standard (if not the standard) employment of 'or' is in the specification of possibilities (one of which is supposed by the speaker to be realized, although he does not know which one).

Grice, Indicative Conditionals, p. 68

- Questions are modeled as sets of propositions...
  {it's raining, it's not raining}
- ...so are disjunctions in many modern approaches to free choice, counterfactuals, exceptional scope-taking, among others
  {John is at home, Mary is at work}
- Inquisitive Semantics: disjunctions are at the core of inquisitiveness — disjunctions are the building blocks of questions

Hamblin (1958), Kratzer & Shimoyama (2002), Fine (2012)

An argument from morphology

A(n in)famous fact

Very many natural languages have the same morphemes for the interrogative complementizer and disjunction operator (and indefinites, more on which later)

- Malayalam is a good example (Jayaseelan, 2004)

John-oo Bill-oo wannu.	Mary wannu-oo?
John-or Bill-or came.	Mary came-or
"John or Bill came"	"Did Mary come"

(cf. also Japanese 'ka', Korean 'na', several variations of Slavic 'li', Polish 'czy', and so on)

Illusory inferences on the erotetic theory

P_1: John is watching TV and Mary is playing tennis, or Bill is doing homework.
P_2: John is watching TV.
C: Mary is playing tennis.

Illusory inferences on the erotetic theory

P_1: John is watching TV and Mary is playing tennis, or Bill is doing homework.
P_2: John is watching TV.
C: Mary is playing tennis.

Question

Are we in a John-watching-TV and Mary-playing-tennis situation, or in a Bill-doing-homework situation?

Illusory inferences on the erotetic theory

P_1: John is watching TV and Mary is playing tennis, or Bill is doing homework.
P_2: John is watching TV.
C: Mary is playing tennis.

Question

Are we in a John-watching-TV and Mary-playing-tennis situation, or in a Bill-doing-homework situation?

Incomplete answer

We are in a John-watching-TV situation.

Illusory inferences on the erotetic theory

P_1: John is watching TV and Mary is playing tennis, or Bill is doing homework.
P_2: John is watching TV.
C: Mary is playing tennis.

Question

Are we in a John-watching-TV and Mary-playing-tennis situation, or in a Bill-doing-homework situation?

Incomplete answer

We are in a John-watching-TV situation.

Jumping to conclusions

I see, so the first answer to the question is the true answer — content overlap

Reasoning and interpretation

Just a scalar implicature

P_1: John speaks English and Mary speaks French, or else Bill speaks German.
P_2: John speaks English.
C: Mary speaks French.

Premise 1 of the illusory inference is interpreted as

Either John speaks English and Mary speaks French and nothing else that is relevant is true or Bill speaks German and nothing else that is relevant is true.

- Illusory inference from disjunction, schematically:
  P_1: $(a \wedge b) \vee c$
  P_2: a
  Conclusion: b
- Strengthened meaning, schematically:
  $P_1{}^+$: $(a \wedge b \wedge \neg c) \vee (c \wedge \neg a \wedge \neg b)$
  $P_2{}^+$: a
  Conclusion: b

Scalar implicature — disjunction

(1) a. John or Mary came to the party.
b. John or Mary, but not both, came to the party.

- Sentence (1-a) seems to mean the same as sentence (1-b). Does this mean that disjunction in natural language is exclusive?

Three possibilities

1. Disjunction in NL is exclusive
2. Disjunction in NL is ambiguous between inclusive and exclusive interpretations
3. Disjunction in NL is semantically {inclusive, exclusive}, the other reading is reached via some process that goes beyond semantics

'Or' is not exclusive

Argument from contradiction

(2) a. John or Mary came to the party. In fact, they both did.
b. John or Mary, but not both, came to the party. In fact, they both did.

The discourse in (2-b) is contradictory, but (2-a) is not. So the first sentences of the two discourses cannot mean the same thing.

Argument without contradictions

(3) a. If John or Mary came to the party, then Bill was pleased.
b. If John or Mary, but not both, came to the party, then Bill was pleased

The two sentences in (3) have different entailments under the supposition that "John and Mary came to the party," so the two 'if' clauses cannot mean the same thing.

'Or' is (probably) not ambiguous

1. If English 'or' were ambiguous between inclusive and exclusive interpretations, we should be able to find languages that have different, unambiguous lexical items for the two meanings. There are some purported examples of this, but they are highly contested.
2. The "ambiguous" sentences in the previous slide should have readings that are contradictory, respectively false.
3. Ambiguity seems not to allow shifting of meanings under ellipsis. But inclusive/exclusive 'or' does. (This argument is actually quite involved, below I give just an illustration.)

(4) John went to the bank$_i$, and Mary did [go to the bank$_{i/*j}$] too.

(5) John saw Mary or$_i$ Bill, and Sue did [see Mary or$_{i/j}$ Bill] too.

Maxim of Quantity

- 'Or' is inclusive, but pragmatic processes (or special semantic processes strengthen its interpretation, making it exclusive in many if not most instances of use.

Maxim of Quantity (Paul Grice)

Make your contribution as informative as needed, no more, no less.

One important consequence of this maxim is that speakers will typically be assumed to be imparting as much information as they can:

Most professors have a life outside of USiena.

Implicature:

Maxim of Quantity

- 'Or' is inclusive, but pragmatic processes (or special semantic processes strengthen its interpretation, making it exclusive in many if not most instances of use.

Maxim of Quantity (Paul Grice)

Make your contribution as informative as needed, no more, no less.

One important consequence of this maxim is that speakers will typically be assumed to be imparting as much information as they can:

Most professors have a life outside of USiena.

Implicature: Some professors do not have a personal life.

If the speaker meant all professors have a personal life, they would have said so.

Problem with a naive Gricean account

(6) a. Some mammals fly.
b. All mammals fly.

The reasoning

(6-b) $\models$ (6-a), so if the speaker thought that (6-b) was true, they would have said so, as (6-b) is more informative than (6-a). Conclusion: the speaker doesn't think that (6-b) is true.

Problem with a naive Gricean account

(6) a. Some mammals fly.
b. All mammals fly.
c. Some but not all mammals fly.

The reasoning

(6-b) $\models$ (6-a), so if the speaker thought that (6-b) was true, they would have said so, as (6-b) is more informative than (6-a). Conclusion: the speaker doesn't think that (6-b) is true.

What about this reasoning?

(6-c) entails (6-a), so if the speaker thought that (6-c) was true, they would have said so, as (6-c) is more informative than (6-a). Conclusion: the speaker doesn't think that (6-c) is true.

Basic ingredients of a good account of SIs

(7) a. some
b. all
c. some but not all

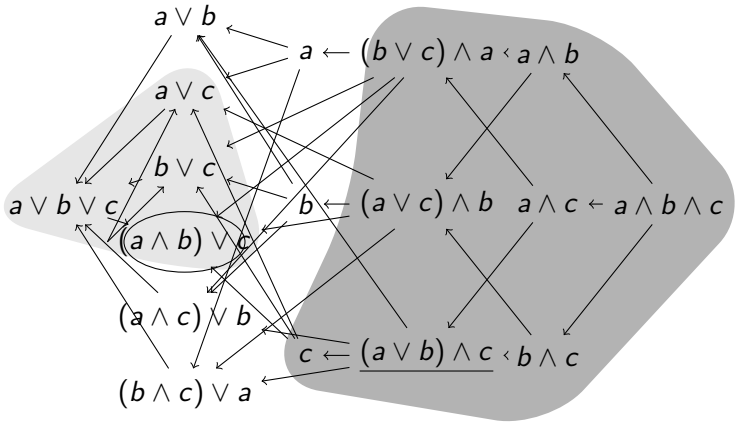
- A notion of an alternative set, such that (7-b) is an alternative to (7-a), but (7-c) is not.
- A selection procedure that picks the right alternatives from the alternative set, so as not to contradict the literal meaning of the sentence.
- An explanation of the epistemic step: Gricean reasoning allows us to conclude from "speaker didn't say X," under the right circumstances, that "speaker doesn't believe X," but what gets us to "speaker believes $\neg X$"?

Calculating scalar implicatures

1. Compute the alternatives to S that are at most as complex as S (Katzir, 2007). These are the things that the speaker could have said but chose not to.
2. Collect those alternatives S' that are (1) alternatives to S and (2) strictly stronger than S. Call this set A.
3. Compute primary implicatures: for each sentence $S' \in A$, "the speaker does not believe that S'."
4. Compute secondary implicatures: for each $S' \in A$ such that the negation of S' does not contradict the literal meaning of S or any of the primary implicatures of S, conclude (that the speaker believes) that S' is false.
5. Call the conjunction of the literal meaning of S together with all of its secondary implicatures the strengthened (exhaustive) meaning of S.

What the speaker could have said

These are all the alternatives for a sentence of the form $(a \wedge b) \vee c$:



Getting the implicature

Alternatives that will give rise to secondary implicatures:

$\{\neg((a \vee b) \wedge c), \neg(a \wedge c), \neg(b \wedge c), \neg(a \wedge b \wedge c)\}$

Equivalently:

$(\neg a \wedge \neg b) \vee \neg c$

Conjoined with the literal meaning:

$((a \wedge b) \wedge \underline{((\neg a \wedge \neg b) \vee \neg c)}) \quad \vee \quad (c \wedge \underline{((\neg a \wedge \neg b) \vee \neg c)})$

Equivalently:

$(a \wedge b \wedge \neg c) \vee (c \wedge \neg a \wedge \neg b)$

Too many alternatives

- Every theory of scalar implicature needs to specify what the relevant alternatives are.
- But most proposals for alternative-set generation in the literature involve rapidly growing sets as a function of the number of atoms in the input.

	2	3	4	n
1. Propositions	16	256	65,536	$2^{(2^n)}$
2. Positive propositions	4	18	166	Dedekind numbers: $M(n) - 2$
3. Katzir (2007)	20	552	20,679	$\sum_{k<n}(2n-1)^{k+1}2^k - k$

Table 1: Number of alternatives by procedure, for a source with 2, 3, 4, and n atoms.

Just a scalar implicature?

- Illusory inferences can be generated with quantifiers doing the job of conjunction and disjunction

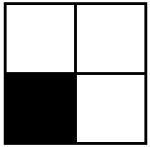
Mary has met every king or every queen of Europe.
Mary has met the king of the Netherlands.
Mary has met the king of Spain.

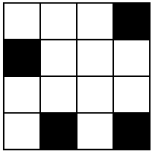
Some pilot writes poems.
John is a pilot.
John writes poems.

- Universals get acceptance rates around 85%, same as the propositional case, but indefinites are at around 40%
- There's an implicature account of IIFD with universals:
  Mary has met every king and no queen of Europe, or else she has met every queen and no king.
- But not with indefinites: ? There is exactly one pilot and she writes poems.

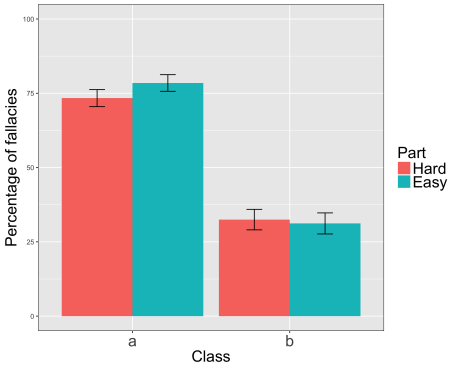
Fewer fallacies under cognitive load (with Léo Picat)

- People are less likely to compute scalar implicatures under cognitive load (de Neys & Schaeken, 2007). We showed these results carry over to inference tasks





Results



- Mixed-effects model showed a significant interaction between class of fallacy (implicature vs. no implicature) and difficulty of the memory task
- Only the fallacies where an implicature account exists showed the effect.
- Small effect size, comparable to what was found for truth-value-judgment tasks.

Related work in progress

- Developmental studies: children until around the age of 6 derive fewer scalar implicatures than adults. We predict that they should make fewer mistakes than adults in those fallacies that are due to implicatures.
- Priming studies: it is possible to affect the rate at which adults get scalar implicatures on a given scale by priming them on some other scale. This should carry over to reasoning tasks along the same lines as the dual-task study. But with potentially much larger effect sizes!
- Fallacies in real life: today I presented work on very pure reasoning tasks, but these illusory inferences from disjunction can be recast in contexts with real-life relevance, such as marketing and advertising. We are interested in recruiting methods from psycholinguistics to develop strategies to help people make better decisions.

Conclusions

- Linguistic semantics and the psychology of reasoning display significant overlap in object of study — we need to reach convergence!
- The methods of linguistic analysis are of the utmost importance in our understanding of how human reasoning works — only through careful linguistic analysis can we understand what is a reasoning mistake and what is an interesting interpretive process.
- Reasoning matters greatly. By contributing to our understanding of reasoning, linguistics reaches a new dimension of relevance and utility, both as fundamental science and as a source of inspiration for applications.