WARSAW WORKSHOP
IN PHILOSOPHY OF LANGUAGE

LOGIC AND MEANING

Warsaw, 13–14 May 2016

Book of Abstracts

Contents

Programme

First Day, 13th May 2016

University of Warsaw Library, Dobra 56/66, room 256

9:30-9:40	Inauguration
9:40-10:20	*Indexicals in the Context of Attitude Ascriptions* **Katarzyna Kijania-Placek**
10:20-11:00	*The Content of Perceptual Demonstratives* **Adriana Pavic**
11:10-12:40	*Semantic Entry Points for Speaker's Meaning* **François Recanati**
12:40-14:00	Lunch
14:00-15:30	*Analyticity, Explanation, and the Justification of Basic Logical Laws* **Gillian Russell**
15:40-16:20	*Entitlement and Claiming Knowledge of Validity* **Ben Baker**
16:20-17:00	*Assumptions, Hypotheses, and Antecedents* **Vladan Djordjevic**
17:20-18:00	*Directival Theory of Meaning: From Syntax and Pragmatics to Content* **Pawel Grabarczyk**
18:00-18:40	*Quotation and the General Theory of Communication* **Roman Godlewski**

Second Day, 14th May 2016

University of Warsaw Library, Dobra 56/66, room 256

10:00-10:40	*Are Modal Operators Logical?*
	Owen Griffiths
10:40-11:20	*What Is Natural Logic of Natural Language?*
	Jakub Szymanik, Ivan Titov, Fanghzou Zai
11:30-13:00	*Vagueness, Context and Disagreement*
	Joanna Odrowąż-Sypniewska
13:00-14:30	Lunch
14:30-16:00	*What Psychology of Reasoning Can Tell Us About the Meaning of Indicative Conditionals*
	Karolina Krzyżanowska
16:10-16:50	*Conditionals and Content Connection – An Experimental Study*
	Wojciech Rostworowski, Natalia Pietrulewicz, Marcin Będkowski
16:50-17:30	*Nondeclaratives and Logical Words: Their Semantics, Pragmatics, and Logic*
	Dan Boisvert
17:40-18:20	*Processing Affirmation and Negation in Contexts With Unique or Multiple Alternatives*
	Maria Spychalska, Viviana Haase, Jarmo Kontinen, Markus Werning

Semantic Entry Points for Speaker's Meaning

François Recanati

Institute Jean Nicod, CNRS

Contrary to a widespread idealization, grammatical meaning does not determine assertoric content, but merely constrains it. Speaker's meaning necessarily comes into play. In this talk, I am concerned with the extent of the phenomenon. When and where, exactly, does speaker's meaning come into play in fixing assertoric content ?

Analyticity, Explanation, and the Justification of Basic Logical Laws

Gillian Russell

University of North Carolina at Chapel Hill

Recent work on analyticity distinguishes two kinds, metaphysical and epistemic. I will argue that the distinction allows for a new view in the philosophy of logic according to which the claims of logic are metaphysically analytic and have distinctive modal profiles, even though their epistemology is holist and in many ways rather Quinean. This view combines some of the more attractive aspects of the Carnapian and Quinean approaches to logic, whilst avoiding some famous problems.

Vagueness, Context and Disagreement

Joanna Odrowąż-Sypniewska

University of Warsaw

In my talk I'll focus on faultless disagreement as it appears in disputes involving vague predicates and predicates of personal taste. I'll propose a contextualist conception of vagueness that combines indexical and nonindexical contextualism. It has been argued that the problem with contextualism concerning vague assertions is that while indexical contextualism makes impossible any genuine disagreement concerning ascriptions of vague properties to objects, nonindexical contextualism either makes faultless disagreement concerning borderline cases impossible or else it leads to indexical contextualism. I'll suggest a new account of unidimensional vague predicates, according to which in clear cases "a is F" means "a is F simpliciter", whereas in borderline cases it means "a is F-according-to-me". I'll try also to address two possible objections: that faultless disagreement is spurious on my account and that the idea that assertions concerning borderline and clear cases have different contents is ad hoc.

What Psychology of Reasoning Can Tell Us About the Meaning of Indicative Conditionals

Karolina Krzyżanowska

Munich Center for Mathematical Philosophy, LMU Munich

Sentences like "If Hilary Clinton is running for the president of the US, I ate a banana at least once in my life" strike us as odd, yet, according to the most prominent theories of conditionals, they should be evaluated as true (or highly acceptable), provided that Hilary Clinton is actually running in the elections and I ate a banana. Most philosophers and psychologists of reasoning consider the intuition that antecedents and consequents should be somehow connected to be a pragmatic rather than a semantic phenomenon. However, no one has offered a satisfactory explanation of how pragmatics of conversation is supposed to account for this intuition. The few psychological studies that introduced relevance manipulations in the design do not help to resolve the debate either. I will present a theory of conditionals that takes the connection between conditionals' antecedents and consequent to be a part of the semantics and argue that this theory is compatible with the hitherto collected data on how people reason with indicative conditionals.

Entitlement and Claiming Knowledge of Validity

Ben Baker

University of Pennsylvania

In view of the problems for rational insight accounts of basic logical knowledge, such as our knowledge that modus ponens is valid (MP), Paul Boghossian (2001) has proposed taking seriously the possibility that such knowledge may be inferential. He argues that we have an entitlement to infer according to basic inference rules such as modus ponens, which entitlement does not consist in warrant to believe them valid, and so are in a position to acquire such knowledge by a rule-circular argument.

Crispin Wright (2001) and (2004) objects that such an account cannot accommodate our being in a position to reflectively *claim* knowledge of, or warrant for, the validity of basic logical laws since being in a position to claim is subject to requirements which entail that justifying such claims on the basis of rule-circular arguments is subject to vicious epistemic circularity (at least if validity claims can only be warranted inferentially). Thus if internalistic warrant is reflectively claimable warrant, then a purely inferential account cannot accommodate internalist warrant.

Wright (2004) takes the problems for rational insight and inferential accounts to motivate an alternative account on which we have a non-inferential, non-evidential warrant, entitlement to believe basic logical laws such as MP. (He calls this an entitlement to *trust* rather than to believe but since nothing here rests on the distinction I talk about entitlement to believe.) He then goes on to suggest that given this

entitlement we are after all in a position to acquire justification for, and so claim knowledge of, these propositions on the basis of a rule-circular argument.

In a recent paper Alexander Oldemeier (2013) argues that the same considerations that Wright advances to explain why certain inferences, among them those made in Moorean anti-sceptical arguments, fail to transmit internalistic warrant, taken together with principles to which Wright is committed, entail that a rule-circular argument will likewise fail to transmit internalistic warrant and thus that epistemic upgrading of entitlement to MP to justification and so knowledge of it is not possible.

Specifically he suggests that just as Wright is a conservative about perception his internalism commits him to a parallel thesis concerning deduction:

> (**Conservatism about Deduction**) It is an enabling condition for a deductive inference to transmit internalistic warrant that the subject already possesses a warrant for its validity.

Then in the same way as Wright's conservatism about perception entails that Moorean arguments are instances of what he calls the *information-dependence template* of transmission failure, (Conservatism about Deduction) entails that the same template will apply to rule-circular arguments, so that they too will fail to transmit internalistic warrant.

Oldemeier considers three possible lines of response. The first of these is giving up (Conservatism about Deduction). He makes two objections to this. First, that it would mean giving up too much of his epistemological project, and be inconsistent with his rejection of

Boghossian's account. Secondly, that without this thesis he loses much of the rationale for holding that we have an entitlement to MP.

In reply I argue that the form of internalism underlying Wright's objections to Boghossian's purely inferential account does not commit him to (Conservatism about Deduction), at least not in a sense in which it is a form of conservatism which parallels his conservatism about perception, and thus not in a sense that entails that the information-dependence template applies to a rule-circular argument. Whether or not this is Wright's position I suggest that it is available to someone working within something like Wright's framework.

First, I argue that whilst Wright's conditions on being in a position to claim commit him to the view that in order for a warrant acquired by deductive inference to be claimable (and in that sense internalistic) one must have independent warrant to believe the inference to be valid, he is not thereby committed to the view that this independent warrant is required to rationally ground belief in the conclusion; Wright may consistently hold that possessing such independent warrant is not a condition on acquiring warrant by inference in accordance with a basic inference rule.

Secondly, I argue that this matters because the information-dependence template applies when the conclusion of the inference, I, is a reflectively appreciable presupposition of the premise, P, when believed on the basis of evidence, e, in such a way that it can be seen to be a proposition for which one requires independent warrant as a condition of acquiring internalistic warrant for P on the basis of e. Thus if it is to apply to a rule-circular argument, (Conservatism about Deduction) must be understood as bearing on the warrant acquisition condition for deductive inference. I suggest that there are

good reasons to reject it so understood in the case of basic deductive inference and reason to think that Wright rejects it.

Finally, I consider Oldemeier's second reason to think that Wright needs (Conservatism about Deduction). Given the above this will not do as it stands but I suggest that Oldemeier does highlight a seeming tension in Wright's position. How can the fact that a rule-circular argument for MP involves MP as a presupposition play the role it is required to in explaining our entitlement to believe MP if it does not do so in a way that debars the argument from being a justification? I consider two lines of response. The first appeals to an essential connection between the rationality of inferring according to a rule and being prepared to judge that the rule is valid, perhaps by reference to Moore's Paradox. Aside from any doubts about the alleged connection I suggest that it is perhaps not clear that any belief to which one is committed non-inferentially justified. The second response is that since doubt concerning the validity of basic logical rules would debar one from rationally engaging in cognitive projects that presuppose them, it must be possible responsibly to believe them to be valid independent of any inferential justification if one considers the matter. I suggest that this response may be more promising.

References:

Boghossian, P. (2001). How are Objective Epistemic Reasons Possible? *Philosophical Studies*, 1-40.

Oldemeier, A. (2013). Entitlement and Epistemic Upgrading. *Analytic Philosophy*, 436-446.

Wright, C. (2001). On Basic Logical Knowledge. *Philosophical Studies*, 41-85.

Wright, C. (2004). Intuition, Entitlement, and the Epistemology of Basic Logical Laws. *Dialectica*, 155-175.

Nondeclaratives and Logical Words: Their Semantics, Pragmatics, and Logic

Dan Boisvert

University of North Carolina at Charlotte

Nondeclarative sentences ('Come home', 'Thank you', 'What time is it?') raise enormous yet still underappreciated problems for the most promising theories of meaning, which aim to explain how we understand complex sentences by understanding their parts and the way those parts are combined, and for the most promising logical theories, which aim to explain the logical relations among which stand the sentences of a language.

The reasons nondeclaratives raise such problems are easy to understand. (a) Nondeclaratives constitute a large portion of natural languages. Indeed, they constitute four of the five major and minor syntactic moods (imperative, interrogative, exclamative, and optative). Moreover, (b) nondeclaratives embed widely within complex sentences and stand in various logical relations. E.g.: (1) embeds and contributes its meaning to (2) and (7); (3) embeds and contributes its meaning to (8); sets (1) and (2) and (3) and (4) are intuitively inconsistent (in some important logical sense); inferences from (5) and (7) to (1) and from (6) and (8) to (3) are intuitively valid (again, in some important logical sense).

(1) Come home.

(2) Don't come home.

(3) Thank you for cleaning the room.

(4) Shame on you for cleaning the room.

(5) The streetlights are on.

(6) You cleaned up the room.

(7) If the streetlights are on, come home.

(8) If you cleaned up the room, thank you.

Any adequate theory of meaning must explain how we understand (2) and the mixed-mood sentences (7) and (8) by understanding, in part, (1) and (3) and any adequate logical theory must explain the senses in which (1) and (3) stand in their respective logical relations with other sentences; if not, such theories are radically incomplete. Finally, (c) neither nondeclaratives *nor the more complex sentences into which they can embed* are truth-apt, so they do not have truth conditions. Unfortunately, (d) the most promising theories of meaning are truth-conditional theories, which explain our understanding of words and sentences in terms of what they contribute to the truth conditions of more complex sentences, and the most promising theories of logic are truth-preservation theories, which explain logical relations in terms of the preservation of truth. Consequently, (e) the most promising theories of meaning and logic can explain neither our understanding nor the logical relations of large portions of natural languages containing nondeclaratives. A problem, indeed.

This presentation will include the important ideas in my ongoing manuscript, *Nondeclaratives: Their Semantics, Pragmatics, and Logic*, which, necessarily, also contains the important ideas of what I take to be the correct accounts of the semantics, pragmatics, and logic of natural language connectives 'and', 'or', 'not', and 'if-then'. The semantic theory, which I call "success condition semantics" (SCS), is a type

of fulfillment theory. SCS exploits the close connections among sentence type, speech-act type, and the types of conversational goals we usually have when performing those types of speech acts—and explains our understanding of sentences in terms of the conditions that would make those typical conversational goals successful. Here is the (extremely) rough idea: 'Go home' is an imperative, and the imperative structure is conventional device that allows speakers to (directly) direct an addressee's behavior. But speakers do not typically want to direct people to do any old thing; rather, speakers typically want to direct people to do what is actually directed—that is the purpose of the imperative mood. Thus, an adequate success-conditional theory of meaning for English would explain our understanding of 'Go home' in terms of what an addressee must do to comply with the sentence's conventional use, namely, that the addressee go home. That is, the success condition for an imperative is its *obedience condition*. Likewise would SCS explain our understanding of any sentence of a natural language. The success condition for an interrogative is its *response condition*; the success condition for an exclamative or optative is its *sincerity condition*; the success condition for a declarative remains its *truth condition*.

The implications of this semantic account include the following:

- SCS does not discard truth-conditional semantics, but generalizes its most fruitful insights to provide a unified theory of meaning. SCS provides a theory of meaning not just for the declarative sentences of a language but for all sentences of a language.
- SCS carves conceptual space for a unified semantic account of natural language logical words, such as 'and', 'or', 'not', and 'if-then'. For example, any sentence of the form 'P and Q'—whether 'P' and 'Q' are both declarative, both nondeclarative, or mixed--

is successful in English (relative to a speaker and time) just in case P and Q are both successful in their respective ways.

- SCS is the consistent with the view that the meanings of natural language logical words could have logical counterparts in classical logic at least in the sense that their respective meanings can be summarized by their respective "success" tables.
- SCS is consistent with the view that there is a unitary logic at least in the sense that it can provide a unified account of logical relations in terms of the preservation of success. For example, SCS is consistent with the view that there is no need for a distinct "logical system" for imperatives or prescriptions, since the more general logical system based on the preservation of success will subsume that based on the preservation of truth and that based on the preservation of (say) prescription.

Given the presentation time of thirty minutes, I will conclude by making the following points, about which others might ask during the question period:

- The oddity of some linguistic constructions, such as 'If the earth orbits the sun, then two plus two equals four', and even the ungrammaticality of some linguistic constructions, such as 'If thank you, then you cleaned the room' is plausibly explained pragmatically (so, I would reject the view in Björnnson (2008, 2011)).
- Some apparent counterexamples to a logical system grounded in success conditions are plausibly explained either pragmatically or by failing to remember that SCS does not assign entities as meaning.

Assumptions, Hypotheses, and Antecedents

Vladan Djordjevic

University of Belgrade

The distinction between the notions from the title is about the difference between arguments and conditionals (premises and antecedents) and about a further difference between two kinds of arguments (two kinds of premises – assumptions and hypotheses). The difference is easily made in artificial languages, and we are familiar with it from our first logic courses (although not necessarily under those names, since there is no standard terminology for the distinction). I will argue that there are ordinary language counterparts of the three notions, meaning that there are in ordinary language two kinds of premises whose behavior in the context of reasoning is nicely captured by the formal properties of their artificial language counterparts. The distinction in artificial languages can be defined in syntactic and semantic terms, while the ordinary language definitions include pragmatic notions. My next observation is crucial: assumptions, hypotheses, and antecedents are easy to mix in ordinary language. That is, we often do not distinguish arguments from conditionals, nor two kinds of arguments, and use the same expression 'if... then...' for all of them. This, I will argue, can lead to mistakes. I will use the distinction to explain away some well known problems and paradoxes: the direct argument, a standard argument for fatalism, McGee's counterexample to modus ponens, a problem with the Ramsey test (Ramsey+Moore=Good), the poisoned dart, the miner's paradox, and a recent counterexample to modus tollens. (In a half an

hour talk there will be time for three or four problems.) These paradoxes are stated in terms of indicative conditionals, but can be easily stated in terms of counterfactuals. In both cases my solution is the same.

Consider:

1. $$\frac{P_1, P_2, \ldots P_n}{C}$$

2. $$\frac{\vdash P_1, \vdash P_2, \ldots \vdash P_n}{\vdash C}$$

3. $$\frac{\models P_1, \models P_2, \ldots \models P_n}{\models C}$$

4. $\{P_1, P_2, \ldots P_n\} \vdash C$

5. $\{P_1, P_2, \ldots P_n\} \models C$

For simplicity sake, I will restrict my claims only to some logical systems (although they might be applied to many more): classical propositional logic, some modal logics based on it (T and S5), and standard conditional logic, by which I mean these modal systems with the addition of a selection function, used for the familiar truth conditions for conditionals in Stalnaker-Lewis style. 1-5 are familiar from basic logic. They contain 4 different arguments: if the usual meaning of the horizontal line is truth preservance (if whatever occurs above is true, then so is the thing below), then the meanings of 1 and 5 are the same.

This is well-known terminology, but let me point to a possible confusion. We tend to talk about object-language formulae P_1, P_2, ... P_n and C above as premises or conclusions in all of the above arguments, including 2 and 3. However, this is not in accordance with the

usual meaning of the line which leads us to understand premises and conclusions of 2 and 3 as meta-language claims. Thus we have to choose:

9. truth-preserving line and premises/conclusions in meta-language, or

10. validity/theoremhood-preserving line and object-language premises/conclusions.

Let us choose 10.

Definition 1

> An *assumption* is an object-language formula used as a premise in an argument of the form 2 or 3.
>
> A *hypothesis* is an object-language formula used as a premise in an argument of the form 4 or 5.
>
> An *argument from assumptions* (*hypotheses*) has the form of 2 or 3 (4 or 5).

If the form of an argument is determined by the form of its premises and conclusion, then our choice of 10 instead of 9 allows us to talk of some arguments of different kind as having the same form. Let us use a double line for validity/theoremhood-preservance. Then all these arguments can be called modus ponens in virtue of their form:

$$\frac{\models A, \models A \rightarrow C}{\models C}$$

$$\{A, A \rightarrow C\} \models C$$

$$\frac{A, A \rightarrow C}{\quad C \quad}$$

$$\frac{A, A \rightarrow C}{C}$$

This is in accordance with our *informal* practice: we do call all of these modus ponens.

Some forms valid for one kind only:

	necessitation	inference both ways gives equivalence	transitivity	contraposition	constructive dilemma
arguments from hypotheses	×	✓	×	×	✓
arguments from assumptions	✓	×	✓	✓	×

The second rule: 24 follows from 23 but not from 22. Transitivity, contraposition pertain to conditionals, not material implication.

22. $\dfrac{\models A}{\models B}$ and $\dfrac{\models B}{\models A}$

23. $\{A\} \models B$ and $\{B\} \models A$

24. $\models A \equiv B$

Definition 2

> An antecedent of a true indicative (counterfactual) conditional is (would be) sufficient, in the given context, for the consequent.
>
> Hypotheses of a valid argument are sufficient, in any possible context, for the conclusion.
>
> Assumptions of a valid argument are such that their special status is, in any possible context, sufficient for the same status of the conclusion

By 'special status' I do not mean truth (otherwise we lose the distinction). 'Special status' is or implies (some kind of) necessity. Semantic/syntactic examples: validity, theoremhood, necessity. Pragmatic

examples: explicit suppositions, propositions entailed by Stalnakerian context set. The latter have a status of 'temporal necessity'. As long as they are part of common ground, their negation is ruled out of consideration (doesn't count as a possibility).

Source of mistakes: unintentionally using as a premise (or an antecedent) not a proposition but its special status. The former requires an argument from hypotheses (or a conditional), the latter – from assumptions.

The direct argument (DA: $A \vee B$ entails $\neg A \rightarrow B$) seems acceptable but makes conditionals equivalent to material implications. This can be explained away by pointing out that DA as an argument from assumptions is valid, but doesn't imply the equivalence. As an argument from hypotheses it does imply the equivalence, but is not valid.

The main problem with McGee's counterexample is a composite conditional that appears trivially true, even though it has a true antecedent and a false consequent. The problem is confusion of conditionals with arguments from assumptions.

Likewise the other problems are explained away by pointing to a confusion of assumptions with either hypotheses or antecedents.

Quotation and the General Theory of Communication

Roman Godlewski

The Author's intuitive starting point of analysis of quotation included the following theses:

I. Quotation goes on both in speech and in writing.

II. Quotation requires that the quoted material is presented in extenso.

III. Translative quotations are equally good as quotations that preserve the language of the original.

Thus the task was to search for a theory of quotation that fulfils all these claims. The Author has realized that in this aim it is necessary to broaden the common paradigm of linguistic research, and to analyze carefully what an act of communication is. The aim of the presentation is to sketch some new ideas in this domain.

An act of communication includes:

I. The sender's intention to evoke a given content in a given recipient's mind with a given activity in given circumstances,

II. The sender's significant activity,

III. The significant circumstances,

IV. Knowing the significant details (activity and circumstances) by the recipient,

V. The process of interpreting this knowledge by the recipient,

VI. Evocation of the intended content in the recipient's mind (understanding) upon the interpretation.

In order to understand communication the concept of reference must be meant broadly. It may not be limited to linguistic propositional acts only, as Recanati wants (2001, p. 647). Every act of moving (leading) the recipient's attention from one object to another is an act of reference.

In communication you employ significant objects. They are parts of the sender's activity or of the significant surroundings. Some of them refer to other significant objects or to generalities, and the Author calls them representatives. Some significant objects are referred to by other significant objects, and we call them referents. There are possible sequences of reference (tree-like graphs in general). Every such sequence begins with a representative (or a group of representatives) that is not a referent (it is always a piece of the sender's significant behavior; usually an act of speech or pointing) and it ends with a referent (or a group of referents) that is not a representative. For example you can point to an arrow painted on the floor that is the first of a sequence of ten arrows leading to a table. Arrows are intermediate representatives and referents. The sequence of reference has twelve elements, and it starts with your gesture, goes through ten arrows and ends with the table.

Generally an object may refer to another or be an element of a group referring to an object by:

I. Pointing to it,

II. Being its effigy,

III. Being its associative,

IV. Being its symbol or

V. Being a hint.

You can make reference to a generality in a few ways. The object or a group of objects may be:

I. A sample of the generality,

II. An associative: a sample of a generality that the recipient would probably associate with the given one,

III. A symbol of the given generality (a general name).

As a symbol the Author means an object which content (meaning) is established by a custom or convention.

Having generalities indicated you can point to objects with them, e.g. with the expression 'the table'. Having relations indicated you can go from some objects to some other, e.g. with the expression 'John's favorite chair'. And also you can go from some generalities to some other, that is, you can combine general concepts and make definitions, like 'red apple'. Lots of other kinds of reference you can build upon these basic acts or others.

A sign is defined as a representative, that is, an object that refers to another one.

The crucial observation is that reference may exist only in an act of communication, and that is a whole whose all the parts are necessary and lack of one of them makes that there is no communication, and hence no reference, and no signs.

This property of signs means that you cannot repeat a sign without performing the relevant communicational act. Say John said to Mary 'I love you' in order to let her know his love to her. If Peter wants to repeat John's words when talking to Mark it would be necessary that he says the words 'I love you' in order to let Mark know his love to

him. Only then it would be a true repetition of the sentence. It requires that all the important details of the communicational act it is used in are reproduced. In case of the sentence 'I love you' it is necessary that the speaker informs the recipient that he loves him. Without that there is no true repetition of the sentence. Similarly when you produce a leg you cannot do it without the rest of the body, you cannot produce a cousin without other members of the family, and you cannot make tip of a stick without the stick. And this is exactly what you do not do when you quote! (See Clark and Gerrig 1990, p. 764, 800, for similar remarks).

This means simply that quotation of a sign employs not the sign but merely the shape of it.

An oral sound or a patch drawn on paper may be employed to refer to many various things. In the context of quotation marks there are three general possibilities:

I. It may refer to a similar physical object and then we call it an **imitation**.

II. It may refer to a sign and then we call it a **citation**.

III. It may refer to the content of the sign and then we call it **quotation**.

Quotation is an act of employing in extenso (by performance or reference) a sample of the same kind as the object used as the sign in a communicational act (the sample may be even the same object as the one used in the original) in order to refer to the content of the sign.

References:

Clark, Herbert, & Gerrig, Richard, 1990, 'Quotations as demonstrations', *Language*, 66(4): 764–805.

Recanati, F., 2001, 'Open quotation', *Mind*, 110: 637–87.

Directival Theory of Meaning: from Syntax and Pragmatics to Content

Paweł Grabarczyk
University of Lodz

The main idea behind directival theory of meaning is that the meaning of expressions is determined solely by the set of rules of confirmation of sentences that contain these expressions. The confirmation rules (called "directives") are divided into four groups: axiomatic directives (sentences which the user has to accept in any circumstances), deductive directives (sentences the user has to accept should she earlier accepted some other sentences), empirical directives (sentences the user has to accept in a given extra-linguistic circumstances) and imperative directives (sentences which, if accepted, must result in motor action of the user). Roughly speaking, the meaning is then defined as the role the expression plays within a syntactic network of these rules called 'the language matrix'. As mentioned above, the crucial point is that some of these rules, called 'empirical directives', correlate linguistic responses with extra-linguistic internal states of the system (defined as receptor activation patterns and motor actions). The idea originates from early work of (Ajdukiewicz 1978) but has never been developed further. I claim that the theory was forgotten because for many years philosophers didn't have conceptual tools to appreciate its innovative nature.

In the first part of the talk three pillars of the theory are thoroughly explained: (1) The syntactic part of the theory, (2) the extra-linguistic part of empirical and imperative directives and (3) the pragmatic

part understood as the record of normative behavior of the analyzed language community. I show how data coming from all of these sources can be encoded into a 'language matrix' and how the matrix generates linguistic content from these data.

In the second part of the paper I present several problems the theory faces and outline ways of solving them. Specifically I address the counterexample presented by Alfred Tarski which shows that the meaning defined by the directival theory cannot fix the reference of terms as two terms the theory interprets as synonymous could still refer to two different objects. Tarski considered an extremely simple language of first order logic (with identity) and enriched it with two directives: A≠B and B≠A, where A and B are extra-logical constants figuring only in these two directives. He then showed that the definitions of the directival theory cause the terms to be synonymous although they clearly cannot refer to the same object. This contrasts with the assumption that the identity of reference follows from the identity of meaning.

I suggest that the best way to cope with this counterintuitive consequence of the theory is to treat the directival theory of meaning as a theory of narrow linguistic content. I argue that if we understand the theory this way and supplement it with some additional tools we can transform it into a modern functional role semantics capable of reducing linguistic narrow content to a combination of syntax and pragmatics. Thus, the modernized directival theory of meaning can be seen as a new argument against radical meaning externalism (Recanati 1994.)

Last but not least I show that the directival theory does not have some of the recognized flaws of competing functional (or conceptual) role semantics (e.g. Block 1986): it makes it easy to convey the

notion of normativity in language (Greenberg, Harman 2007) and sidesteps the Fodor-Lepore dilemma (Fodor, Lepore 1991). I focus on the latter and show that the directival theory is immune to it due to its several unique features of the. The most important of these features is that the theory can be understood as a middle ground between semantic atomism and semantic holism. The reason for it is that the basis the theory defines meaning on (the set of meaning directives) is significantly smaller than the set of all sentences users accept (thus the theory does not fully determine linguistic behavior of the users.)

References:

Ajdukiewicz, K. (1978). Language and Meaning. (In J. Giedymin (Ed.), The Scientific World-Perspective and Other Essays, 1931-1963. (pp. 35–66). Springer)

Block, N. (1986), "Advertisement for a Semantics for Psychology", Midwest Studies in Philosophy, 10 (1), 615-78.

Fodor, J., Lepore, E., (1991), "Why Meaning (Probably) Isn't Conceptual Role", Mind & Language, Vol.6 Nr 4, 328-343

Greenberg, M., Harman G. (2007), "Conceptual role semantics". In E. LePore, B. Smith (eds.), Oxford Handbook of Philosophy of Language, Oxford: Oxford University Press, pp. 242-256.

Recanati, F. (1994). How narrow is narrow content? Dialectica, 48(3/4), 209–229.

Are Modal Operators Logical?

Owen Griffiths

University of Cambridge

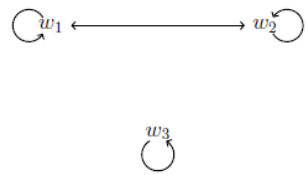
Modal talk is common in natural language. And there are many intuitively valid arguments of Englidh whose validity crucially relies on modal vocabulary, e.g. `It is possible that there is a talking doney; so it is not necessary that there are no talking donkeys'. We can only account for the validity of this argument if we accept modal operators as genuinely *logical*. I will argue that, if we accept a very plausible account of the logical constants – permutation invariance – modal operators should be accepted as logical constants.

The permutation invariance demarcation of the logical constants has much to recommend it: it is philosophically motivated by the thought that logic is topic neutral (see MacFarlane 2000, Ch.6), it can be given precise formulation (see McGee 1996) and all the usual _rst-order logical constants pass the test. Dutilh Novaes (2014), however, has argued that it undergenerates by failing to judge some intu itively logical expressions as logical. Her counterexamples are the S4 and S5 modal operators. I show that her argument is successful against the S4 modal operators but that this should not worry us, since S4 is not the correct logic of logical necessity. The correct logic here is S5 which, I argue, survives her criticisms.

In standard tests for permutation invariance, we consider permutations of the domain of objects, in the modal case, we consider permutations of worlds. In this paper, I will accept this test as the correct way of extending permutation invariance to the modal case, since I will show that even by this test, there is no undergeneration problem.

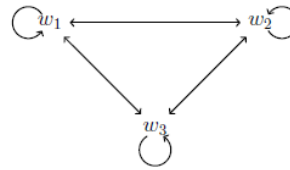
Dutilh Novaes (2014, 93) puts forward a countermodel to the claim that the necessity operator for S4 is permutation invariant. But it is widely acknowledged that the correct modal logic of *logical* necessity as opposed to, say, metaphysical necessity or demonstrability is S5. So we should not be concerned that the S4 necessity operator fails the test. What should trouble the proponent of permutation invariance is the failure of the S5 necssity operator to pass the test.

And in fact the necessity operator for S5 does fail the permutation invariance test, at least when its accessibility relation is an equivalence relation. Consider the following frame:



Let P be true at w_1 and w_2 and let ¬P be true at w_3. Now □P is true at w_1. But, if we permute worlds w_2 and w_3, □P is no longer true at w_1, since it can now access w_3, at which P is false. Hence, the necessity operator for S5 is not permutation invariant when the accessibility relation is an equivalence relation.

This is not sufficient, however, to show that the S5 modal operator is not permutation invariant, for consider the following frame:



Once again, let P be true at w_1 and w_2, and let ¬P be true at w_3. Now, permuting worlds w_2 and w_3 does not disturb the extension of the necessity operator, since P would be false before the permutation. This is because the accessibility is no longer *merely* an equivalence relation but also a universal relation: every world can access every other world.

But S5 can be characterised as having an accessibility relation that is universal. Hughes and Cresswell (1984, 122-3) show that the characterisation of S5 with a universal accessibility relation is equivalent to its characterisation with a relation that is an equivalence relation but not universal. So as long as S5 is characterised as having a universal accessibility relation, it passes this test for permutation invariance, and is therefore logical.

At this point, an objection may be made to the proponent of permutation invariance. Permutation invariance judges S5 modal operators to be logical when interpreted on universal frames, but nonlogical when interpreted on frames that are merely equivalence relations. But the two modal logics are provably equivalent (have the same consequence relation) so surely they should either stand or fall together, with respect to logical nature. Isn't this an embarrassment to the the proponent of permutation invariance?

I will offer two responses to this objection. First, it is crucial here that we are careful when we talk about two operators' being `equivalent'. Permutation in variance is a semantic test for logical nature. It applies to operators semantically understood. It is incoherent, therefore, to claim that the S5 necessity operator with a universal accessibility relation, $\Box_U$, and the S5 necessity operator with an equivalence accessibility relation, $\Box_E$, are the same operator on different presentations. Rather, $\Box_U$ and $\Box_E$ are different operators. Syntactically, they behave the same, but permutation invariance does not apply to syntax.

Second, the proponent of permutation invariance can *explain* why the permutation variance of $\Box_E$ is unproblematic. Permutation invariance is a test for *logical nature* and so, as we have seen, the salient notion of necessity here is *logical* necessity. We are interested what is true at all *logically* possible worlds, as opposed to e.g. physically or metaphysically possible worlds.

The actual world is a logically possible world and the other logically possible worlds are those accessible from the actual world. The logically possible worlds, therefore, will be those in a certain equivalence class of worlds (the equivalence class that features the actual world). All of the other worlds are logically *impossible* worlds. The $\Box_E$ operator is permutation invariant over any particular equivalence class: it is only when we start to permute worlds between equivalence classes that it is permutation *variant*. In particular, then, $\Box_E$ is permutation invariant over the equivalence class of logically possible worlds. And this is all that we should hope for: that $\Box_E$ is insensitive to the particular characters of *logically possible* worlds.

The proponent of permutation invariance can do better, therefore, than merely accept that $\Box_U$ and $\Box_E$ differ with respect to permutation

invariance. They can explain *why* the former is invariant and the latter is variant under permutations of worlds. The latter is *only* invariant when we permute between equivalence classes of worlds. But this is to allow the logically possible worlds to be exchanged with logically impossible worlds, and we are not interested in the latter when we determine the logically necessary truths. Across the equivalence class of logically possible worlds, $\Box_E$ is permutation invariane, which is sufficient.

References:

Dutilh Novaes, C. (2014). The undergeneration of permutation invariance as a criterion for logicality. *Erkenntnis 79*, 81-97.

Hughes, G. and M. Cresswell (1984). *A Companion to Modal Logic*. London: Methuen.

MacFarlane, J. (2000). *What Does it Mean to Say that Logic is Formal?* PhD thesis.

McGee, V. (1996). Logical operations. *Journal of Philosophical Logic 25*, 567-580.

van Benthem, J. (1989). Logical constants across varying types. *Notre Dame Journal of Formal Logic 30*, 315-42.

Indexicals in the Context of Attitude Ascriptions

Katarzyna Kijania-Placek

Jagiellonian University

This talk offers an account of some uses of indexicals in the context of propositional attitude ascriptions, i.e. reports that concern the cognitive relations people bring to bear on propositions. While the contribution of indexicals to the truth conditions of an utterance is usually singular, their interpretation is general in the case of so called descriptive uses. I will propose an interpretation of the descriptive uses of indexicals via a mechanism of descriptive anaphora and apply this mechanism to the case of attitude ascriptions. I will emphasize the role of context both in the suppression of the default referential reading of the indexical, as well as in the reconstruction of the relevant *interpretation* of the whole utterance.

Indexicals are typically considered as vehicles of direct reference. Some contexts of propositional attitude ascriptions make it clear, however, that the singular mode of presentation deployed by an ascriber cannot be attributed to the ascribee. An example has been given by Nunberg in (1993):[1]

(1) The Founders invested me with the sole responsibility for appointing Supreme Court justices. uttered by George H.W. Bush in 1992

[1] In treating (1) as a proposition attitude ascription I assume the propositional analysis of intensional verbs (see Quine 1956; Larson 2002; Moltmann 1997).

Existing accounts of propositional attitudes (for example Quine 1956, 1960; Kaplan 1968; Recanati & Crimmins 1995; Recanati 2000; Richard 1983; Bonomi 1995; Aloni 2005; Kripke 2011) seem to imply that by uttering such a sentence George H.W. Bush might be committed to the absurd claim that the Founders had *de re* thoughts about himself. That is because in most accounts of propositional attitude ascriptions it is assumed that attitude ascriptions that contain indexicals are *de re* ascriptions (see Richard 1983), i. e. such that the mode of presentation of the referent of the indexical does not affect the truth conditions of the belief report and we are usually not told how the subject of the attitude thinks about the referent. My aim in this paper is to offer an account of the reconstruction of the proposition expressed by the original utterance, reported in (1), in terms of a descriptive interpretation of indexicals that would not have such unintuitive consequences.

My thesis is in direct relation to Recanati's writings on propositional attitude ascriptions. Persuaded by the arguments of Morgan (2011), Recanati concedes in (2012) that from the fact that an indexical has been used in the that-clause of an attitude ascription we cannot infer that no mode of presentation of the referent is ascribed to the believer. But, Recanati maintains, we can infer that the ascribed mode of presentation is singular. He thus seems to support the singularity of ascribed believe thesis:

> Singularity of the ascribed belief
>
> An indexical within the that-clause of an attitude ascription indicates a singular mode of presentation of the referent in the ascribed belief.

Contrary to that I hope to be able to show that there are situations in which it is obvious from the context that the reported belief could not have been a singular one. In such cases it seems that the indexical used in the that-cause is just exercised but not attributed, contrary to Recanati's thesis. This would happen when the referent of the indexical is present in the context of the ascription but, for reasons obvious from the context, could not have been present during the reported utterance, like in example (1). In this case it is obvious that the Founders could not have had *de re* thoughts about George H.W. Bush and the hearer, if aware of the fact, does not interpret the president as claiming so much. Additionally, the hearer is able to reconstruct the reported general belief by implicitly relying on the mechanism of descriptive anaphora.

On the descriptive anaphora interpretation, George H.W. Bush (the person, not the name) is the extra-linguistic antecedent of this token of 'he' and points to his silent property of 'being the president of the United States' ('US-president' for short). The quantifier that gives the structure to this general proposition is the binary universal quantifier and the property obtained from the context serves as its context set. As a result, we obtain the following structure of the original declaration (RASCJ is short for 'having been given the responsibility for appointing Supreme Court justices'):

EVERYx(US-PRESIDENT(x); RASCJ(x))

– 'Every president of the United States has been given the responsibility for appointing Supreme Court justices', which seems to be the intended interpretation of the reported belief.

References:

Aloni, M.: A formal treatment of the pragmatics of questions and attitudes. Linguistics and Philosophy 28(5), 505–539 (2005).

Bonomi, A.: Transparency and specificity in intensional contexts. In: P. Leonardi, M.S. (ed.) On Quine. New Essays, pp. 164–185. CUP, Cambridge, MA (1995).

Kaplan, D.: Quantifying in. Synthese 19, 178–214 (1968).

Kripke, S.A.: Unrestricted exportation and some morals for the philosophy. In: Philosophy Troubles, pp. 322–350. OUP, Oxford (2011).

Larson, R.: The grammar of intensionality. In: Logical Form and Language, pp. 228–262. Clarendon Press, Oxford (2002)

Moltmann, F.: Intensional verbs and quantifiers. Natural Language Semantics 5(1), 1–52 (1997).

Morgan, D.: First Person Thinking. Ph.D. thesis, University of Oxford (2011).

Nunberg, G.: Indexicality and deixis. Linguistics and Philosophy 16, 1–43 (1993).

Quine, W.V.O.: Quantifiers and propositional attitudes. The Journal of Philosophy 53, 177–187 (1956).

Quine, W.V.O.: Word and Object. MIT Press, Cambridge, MA (1960).

Recanati, F., Crimmins, M.: Quasi-singular propositions: The semantics of belief reports. Proceedings of the Aristotelian Society, Supplementary Volumes 69, 175–193+195–209 (1995).

Recanati, F.: Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation. The MIT Press, Cambridge, Massachusetts (2000).

Recanati, F.: Mental Files. OUP, Oxford (2012).

Richard, M.: Direct reference and ascriptions of belief. Journal of Philosophical Logic 12, 425–45 (1983).

The Content of Perceptual Demonstratives

Adriana Pavic

Georg-August-Universität Göttingen

According to Tyler Burge "accounting for the cognitive value of demonstratives" is among "the most difficult and persistent specific problems within the philosophy of language" which build pressure toward a shift to the philosophy of mind (Burge 1992:28).

In this talk, I try to get at the bottom of this problem by taking a closer look at the paradigm case of demonstrative reference: bare proper demonstratives used to refer to perceived empirical objects. Starting out from Perry's famous example where someone refers twice to the aircraft carrier Enterprise by pointing to different parts of it (Perry 1977:483), with minor modifications we arrive at a particularly tricky version of Frege's puzzle:

> **(FP$_D$)** This [pointing to the stern of the Enterprise] is the same (ship) as this [pointing to the bow of the Enterprise].

We cannot account for the cognitive value of assertions like (FP$_D$) by appealing to the Kaplanian character of the expressions used, or their 'roles' as Perry (1977) has it, because token of the same linguistic type are involved in the assertion. The aim of this talk is to provide a Neo-Fregean account of the content of perceptual demonstratives capable of dealing with the problem posed by assertions like (FP$_D$).

In the first part of my talk, I introduce Künne's (1982, 1992) theory of Hybrid Proper Names as an alternative to Neo-Russellian concep-

tions, endorsing Textor's (2007) version of it with minor modifications. According to the proposed version of the theory of Hybrid Proper Names, it is not the token of the demonstrative expression by itself which refers to a given object, but instead we have to conceive of the token demonstrative in combination with the accompanying demonstration as *a hybrid referring expression*. Treating demonstrations as non-linguistic parts of a hybrid referring expression has considerable advantages over standard views considering them as 'context parameters'.

Though superior to recent Neo-Russellian conceptions like Salmon's (2002), which appeals to demonstrations to try to account for the cognitive value of assertions like (FPD) in terms of *the cognitive way to the content* asserted instead of the content itself, the theory of Hybrid Propoer Names does not say anything about the nature of the senses expressed by tuples of token demonstratives and demonstrations, so in my view it does not go all the way down in explaining the cognitive value of assertions like (FPD).

In the second part of the talk, I set out to motivate non-descriptive, object-dependent Fregean senses. Despite obviously being considered as outworn by many (Bach 2010), I show that the idea of *de re-senses* dating back to Evans (1982) and McDowell (1984) once elaborated has an enormous explanatory power.

A sense expressed by a hybrid referring expression composed of a linguisitc token and a demonstration cannot have the form of a definite description, because no such description is capable of accounting for it's cognitive value. But as Evans (1981, 1982) and McDowell (1984) have been at pains to emphasize the Fregean should not be forced into this "descriptive corset" anyway. To identify the referent and grasp the sense expressed by utterances like (FPD), linguistically

encoded information has to be integrated with non-linguisitic (perceptual) information. I develop an outline of a new account of de re sense which I consider to be a suitable extension of the theory of hybrid proper names. This involves addressing the following important questions:

- Should perceptual contents just take over the role of a Fregean sense?
- How can we individuate de re senses, what are their identity-conditions?
- How can we guarantee the rigidity of demonstrative reference within such an account?
- How can a de re sense uniquely determine an object?
- In what sense are de re senses objective and intersubjectively accessible?

The upshot is that the Neo-Fregean theory I sketch captures the essential insight that qua human beings our access to objects is necessarily perspectival while at the same time holding the fundamental insights of theories of direct reference (Kaplan 1989a, Recanati 1993) in place.

In the last part of the talk, I situate the outcome within the broader debate concering the semantics/pragmatics-boundary.

Though it's determination is *linguistically controlled*, the content of bare proper demonstratives cannot be *linguistically determined*, hence I argue that a Semantic Minimalist should give up propositionalism (contra Borg 2004) while retaining anti-intentionalism as a central minimalist tenet. In this way, we can retain Borg's account of sentences like

(S) This is red.

in terms of conditionalised truth-conditions along the following lines:

> **(STC)** If the speaker of 'this is red' refers with the utterance of 'this' therein to α and to nothing else, then that utterance is true if and only if α is red.

while accepting that the propositions expressed with sentences like (S) cannot be derived from such abstract schematic contents. Thus, we need to hold apart the (non-propositional) truth-conditions of sentences like (S) and the propositions expressed by utterances of such sentences.

The characteristic features of paradigmatic bare demonstratives like "this" and "that" – almost entirely lacking descriptive content and standing in need of an accompanying demonstration – force us to take into account both the directing intentions of the speaker (Kaplan 1989b, Perry 2009) and the hearer's perceptual acquaintance with the intended referent which are necessary to express and respectively grasp the content of a demonstrative. Thus we find ourselves exactly in the border area between philosophy of language and philosophy of mind Burge is taking about.

Conditionals and Content Connection – An Experimental Study

Wojciech Rostworowski, Natalia Pietrulewicz, Marcin Będkowski
University of Warsaw

The aim of this talk is to present an experimental study on semantics of conditionals. A felicitous statement of a conditional, i.e., sentence of the form "If *p*, then *q*", conveys that there is a content connection between *p* and *q*. Generally, this content connection has an implicative character, that is to say, it assumes that *p* somehow implies *q*; however, depending on the context, this connection may be of a very different kind. In particular, a conditional utterance may convey that *p* is a *cause* of *q*, *p* is *evidence* of *q*, *q* is *conclusion* from *p* etc. Various theories of conditionals agree that the content connection between the antecedent and the consequent is not significant to truth conditions of conditionals.[2] Namely, a semantic evaluation of "If *p*, then *q*" (true/false) does not have to include the fact whether there actually obtains any particular connection between *p* and *q*. A consequence of this view is that the conditionals with sub-clauses which apparently lack any connection (such as conditions (1) and (2) below) may be evaluated as true/highly credible, despite being very odd and presumably unacceptable by a normal user of language:

(1) If two plus two equals four, then Warsaw is the capital of Poland.

(2) If Paris is the capital of France, then the Earth orbits the sun.

[2] Including Bennett (2003), Edgington (2007), Grice (1989), Stalnaker (1968), Thomson (1990).

Indeed, most theories of conditionals assume that a conditional with a true antecedent and a true consequent is true. The oddity of (1) and (2) is usually explained by an appeal to Grice's theory of communication (cf. Bennett 2003: 116, Edgington 2007: 158-160, Thomson 1990.) Roughly, by making an utterance of "If *p*, then *q*", a speaker implicates – in virtue of Maxims of Quality and Quantity – that she believes in a substantial connection between *p* and *q* (and not merely that, e.g., both sentences are true.) In other words, the content connection arises as a conversational implicature and thus (1) and (2) are misleading because of having false implicatures. (Let us call this view IC.)[3]

In our talk, we want to examine whether the content connection in conditionals is actually nonsignificant to their truth conditions and how it should be accounted for – in particular, whether IC is correct. Our examination employs experimental methods in an essential way. We present two experiments on conditionals. The first experiment examines the reaction of ordinary speakers to such conditionals as (1) and (2). In line with the predictions of IC, users of language are inclined to evaluate such conditionals as "true but odd" rather than "false" or "nonsensical". The second experiment applies a test from the so-called "reinforceability" designed to establish whether a given content P associated with a sentence-type S is a part of a semantic content of S or rather an implicature of S (see Sadock 1979, Horn 1991.) The results of this test indicate that a specific content connection expressed by a conditional in a particular context is a part of what is *expressed* rather than what is *implicated*.

[3] On the other hand, Björnsson (2008, 2011) argues that a particular content connection between the antecedent and the consequent is expressed by a conditional in a given context and not merely implicated.

Based on the outcome of our both experimental studies, we propose an account of conditionals, according to which, conditionals have some literal truth/acceptance conditions which do not imply any content connection between the antecedent and the consequent. However, in a context, the truth conditions of a conditional utterance may be more complex in the sense that they imply a particular content connection between the constituent-clauses. The process of truth-conditional enrichment is akin to modulation rather than saturation (cf. Recanati 2010.)

References:

Bennett, Jonathan, (2003), *A Philosophical Guide to Conditionals*, Oxford: Clarendon Press.

Björnsson, Gunnar, (2008), "Strawson on 'If' and Hook", *South African Journal of Philosophy* 27(3): 24-35

Björnsson, Gunnar, (2011), "Towards a Radically Pragmatic Theory of If-Conditionals", in K. Turner (ed.) *Current Research in the Pragmatics/Semantics Interface. Making Semantics Pragmatics*, vol. 24., Brill Academic Publishers.

Edgington, Dorothy, (2007), "On Conditionals" in D.M. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd Edition, Volume 14: 127–221.

Grice, Herber Paul, (1989), *Studies in the Way of Word*, Cambridge: Harvard University Press.

Horn, Laurence, (1991), "Given as New: When Redundant Affirmation isn't", *Journal of Pragmatics* 15, 313-336.

Recanati, François, (2010), *Truth-conditional Pragmatics*, Oxford: Oxford University Press.

Sadock, J. M. (1979), “On Testing for Conversational Implicature”, in *Syntax and Semantics*, vol. 9 *Pragmatics*, (ed. By P. Cole), New York: Academic Press.

Stalnaker, Robert, (1968) “A Theory of Conditionals” in *Studies in Logical Theory*, *American Philosophical Quarterly* (Monograph Series, 2), Oxford: Blackwell, pp. 98–112.

Processing Affirmation and Negation in Contexts with Unique or Multiple Alternatives

Maria Spychalska, Viviana Haase, Jarmo Kontinen, Markus Werning
Ruhr University of Bochum

Negative sentences have been claimed to be more complex than their affirmative counterparts, which has impact on the comprehension process. Regarding the cognitive processing of negative sentences it has been observed that sentences containing a negation are harder to process (e.g. Just & Carpenter, 1971; Carpenter, Just, Keller, Eddy, Thulborn, 1999; Lüdtke, Friedrich, de Filippis, Kaup, 2008), resulting in longer reading times, higher error rates and longer reaction times. This observation raises the question of how negated concepts are represented and how negation is integrated into the sentence meaning. The *Two-Step-Simulation-Hypotheses* by Kaup, Lüdtke & Zwaan (2006) predicts that negation is integrated into the sentence meaning in two steps. At the first step, the negated state of affairs is simulated by the comprehender and the actual state of affairs is simulated only at the second step. Hence, according to Kaup and colleagues, in order to process a sentence such as *The door is not open* one needs to simulate first the open door before simulating the actual situation (closed door). However, the case of polar adjectives that have clear opposites such as *closed* vs. *open* is a special one, since they allow for the identification of the negative of a predicate (e.g. *not open*) through its affirmative opposite (closed). Compared to cases where there is no clear opposite (e.g. *red*, *triangular*) the processing of the negation of polar adjectives can be facilitated. Indeed it has been shown in an

Eye-Tracking-Experiment by Orenes, Beltrán & Santamaría (2014) that the presence of one affirmative alternative facilitates the processing of negative sentences compared to cases where there are several affirmative alternatives.

In our experiment we addressed the question of whether the processing of negation is facilitated in those cases with a unique referent relative to the cases with multiple alternative referents. We ran our study using the method of event-related potentials, which has a high temporal resolution and therefore is frequently used for the investigation of sentence processing. The experiment had the form of a sentence-picture-verification paradigm. We used a 2 x 2 design with the factors: (i) **context model** (unique vs. multiple referent) and (ii) **polarity** of the target sentence (affirmative vs. negative). First the pictures were presented creating the context model and afterwards the target sentence was presented word-by-word on a screen while the EEG was recorded. The pictures depicted three different objects (all of same gender)[4] out of which either one or two were then marked with a frame of red or green color. A green frame was used to indicate that an object was chosen and hence the unframed one(s) is/are unchosen by a virtual agent. A red frame was used to indicate that an object was not chosen and hence the unframed one(s) is/are chosen. Beforehand our subjects were informed that in the following they will observe a person's moves in a game where she or he will select or unselect objects. They were informed about the meaning of the different frames accordingly.

[4] The experiment was done using German sentences and with German natives speakers.

The target sentence always stated which object was chosen or unchosen and always referred to an **unframed** object[5]. Therefore, framing two out of three objects leaves only one possible and therefore **unique** referent (unframed picture) for the object named in the target sentence whereas framing only one object leaves two and hence **multiple** possible referents (unframed pictures) for the object named in the target sentence. The target sentence expressed which object was chosen (**affirmative**) or was not chosen (**negative**), e.g. *Julia hat nicht die Pflaume ausgewählt* (*Julia has not chosen the plum*). The other two conditions accordingly used one (multiple) or two (unique) green frames and the negative version of the sentence, e.g. *Julia hat nicht die Pflaume ausgewählt* (*Julia has not chosen the plum*). All stimuli sentences gave true information with respect to the pictures. Subjects had to respond by clicking a button whether the sentence was true or false with respect to the previously seen pictures. To balance out the material and the required responses we added false fillers.

Example conditions:

Affirmative Unique

Target Sentences (critical word underlined):

Julia hat die Pflaume ausgewählt.

(*Julia has chosen the plum.*)

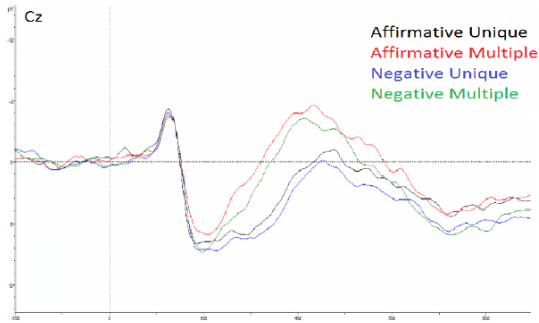
Affirmative Multiple

Julia hat die Pflaume ausgewählt.

(*Julia has chosen the plum.*)

[5] In a second study we reversed the frames (meaning that the sentence always referred to a framed object) to test whether this affects the results. Everything else was kept identical. For brevity we only present results of the original framing in this abstract. In the experiment using reversed frames we observe very similar results.

The Picture shows the Grand Averages (N=20) at the position of the critical word for all four conditions. The preliminary analysis indicates a clear N400 effect for the multiple conditions compared to the unique ones, which is independent of the sentence polarity. This effect indicates that the processing of the unique referent is facilitated relative to the case when there are multiple referents available in the context model. Whereas the *Two Step Simulation Hypothesis* predicts that the contrast between multiple and unique cases should be larger for negative than affirmative sentences, our study does not support this claim.



Furthermore, this result supports the view that the N400 is inversely correlated with the expectancy of the critical word in a context.

Additionally, we observed a late positivity effect for the negative compared to the affirmative conditions, which is in line with the claim that negation is harder to process than affirmation. Therefore, our study provides an interesting insight into the question of how negative sentences are processed.

References:

Carpenter, P.A., Just, M.A., Keller, T.A., Eddy, W.F., Thulborn, K.R. (1999): Time course of fMRI-activation in language and spatial networks during sentences comprehension. *Neuroimage* 10. 216-224.

Kaup, B., Lüdtke, J., Zwaan, R.A. (2006): Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed?, *Journal of Pragmatics* 38. 1033-1050.

Lüdtke, J., Friedrich, C.K.De Filippis, M., Kaup, B. (2008): Event-related potentials correlates of negation in a sentence-picture verification paradigm, *Journal of Cognitive Neuroscience* 20:8. 1355-1370.

Orenes, Beltrán & Santamaría (2014): How negation is understood: evidence from the visual world paradigm. *Journal of Memory and Language*. 74. 36- 45.

What is Natural Logic of Natural Language?

Jakub Szymanik, Ivan Titov, Fanghzou Zai

University of Amsterdam

Natural language contains an abundance of reasoning patterns. Historically, there have been many attempts to capture their rational usage in normative systems of logical rules. However, empirical studies have repeatedly shown that human inference differs from what is characterized by logical validity. In order to better characterize the patterns of human reasoning, psychologists have proposed a number of theories of reasoning. In this paper, we combine logical and psychological perspectives on human reasoning. We develop a framework integrating Natural Logic and Mental Logic traditions. We model inference as a stochastic process where the reasoner arrives at a conclusion following a sequence of applications of inference steps (both logical rules and heuristic guesses). We estimate our model (i.e. assign weights to all possible inference rules) on a dataset of human syllogistic inference while treating the derivations as latent variables in our model. The computational model is accurate in predicting human conclusions on unseen test data (95% correct predictions) and outperforms other previous theories. We further discuss the psychological plausibility of the model and the possibilities of extending the model to cover larger fragments of natural language.

The model It is based on a natural logic proof system by Geurts [3]. Similar to Rips' [7] proposal, we take the set of syllogistic sentences as the mental representation of reasoning. Namely, the reasoner maintains a set of sentences in the working memory to represent the

state of reasoning, or more specically, the reasoner keeps a record of the sentences that he considers true at the moment. We will refer to each representation as a **state**. Reasoning operations change the mental states. When performing reasoning, the reasoner generates a sequence of states in the working memory, where the initial state is the set of premises, and the nal state contains the conclusion. These states are linked by the **reasoning events**, which can be a specic adoption of an inference rule.

We formulate a generative probabilistic model of reasoning. First, reasoners conduct formal inferences, adopting possible logical rules with different probabilities (related to the cognitive diculty of the rule or some sort of reasoning preference). Each inference rule, $r \in R$ is adopted with a different probability specied by the associated weight w_r (*a tendency parameter*) which is estimated from the data. Additionally, the reasoner may adapt a guessing scenario, e.g., in case of a very complex inference. When the reasoner enters the guessing scenario, the probability that the reasoner guesses `nothing follows' is negatively correlated with the **informativeness** level (see [2]) of the premises, i.e., the amount of information that the premises carries: the more informative the premise, the less faith the reasoner have for a 'nothing follows' conclusion. The reasoner chooses the remaining options with probabilities determined according to the atmosphere hypothesis. This hypothesis proposes that a conclusion should t the premises' `atmosphere', namely, the sentence types of the premises [1]. In particular, whenever at least one premise is negative, the most likely conclusion should be negative; whenever at least one premise contains `some', the most likely conclusion should contain `some' as well; otherwise the conclusion are likely to be armative and universal. The probability that a subject could arrive at a

particular syllogistic conclusion is estimated from the tree by summing over all the leaf nodes containing the conclusion. Consequently, we can obtain posterior distribution of conclusions given the premises. These posterior distributions (for each premises) can be treated as model predictions, and we evaluate them (on unseen test set) against the distribution of human conclusions.

Results We use the data from the meta-analysis by Chater and Oaksford [2]. We randomly select 50% of the premises (i.e., half of the dataset) and use the corresponding examples as the training data. The rest of the data is used for evaluation. We use maximum likelihood estimation to obtain the parameter values.

We mainly use the evaluation method proposed in [5], which is based on the signal detection theory. The authors assume that the conclusions of the participants are noisy, that is unsystematic errors occur frequently. Hence, they classify the experimental data into two categories: those conclusions that appear reliably more often than chance level, which a theory of the syllogisms should predict to occur; and those that do not occur reliably more than chance level, which a theory should predict will not occur. In our context, there are five possible conclusions that can be drawn by subject. The chance level is thus 20%. We count a conclusion as reliable if it is drawn signicantly often, i.e., in at least 30% of the trials. The model is doing a good job, its proportion of correct predictions approximating a 95%.

Conclusions We have developed a preliminary framework of combining Natural Logic and data-driven inference weights and applied it to model syllogistic reasoning. The computational model learns from the experimental data, and as a result it may represent individual differences and explains subjects' systematic mistakes. This is

achieved by assigning weights to all possible inference rules using machine-learning techniques and available data. This simple setting solves the **logical omniscience problem**: not all derivations are available. Moreover, the approach takes into account various cognitive factors. For instance, the model enables the agents to adopt illicit conversions (e.g., yielding `All A are B' from `All B are A') in order to explain some systematic errors.

The syllogistic fragment is an informative yet small arena for theories of reasoning. A natural next step would be to extend the model to cover a broader fragment of natural language by exploring existing Natural Logics [4] and designing new logics. The Natural Logics are usually computationally very cheap [6]. This guarantees that our models will easily scale-up to natural language reasoning. The computational complexity analysis will allow assessing the resources and strategies required to perform the reasoning tasks, cf. [8]. This in turn should open new ways of comparing our approach with other frameworks in psychology of reasoning.

References:

[1] Ian Begg and J Peter Denny. Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning errors. *Journal of Experimental Psychology*, 81(2):351, 1969.

[2] Nick Chater and Mike Oaksford. The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38(2):19-258, 1999.

[3] Bart Geurts. Reasoning with quantiffiers. *Cognition*, 86(3):223-251, 2003.

[4] Thomas Icard III and Lawrence Moss. Recent progress in monotonicity. *LiLT (Linguistic Issues in Language Technology)*, 9, 2014.

[5] Sangeet Khemlani and PN Johnson-Laird. Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3):427, 2012.

[6] Ian Pratt-Hartmann. Fragments of language. *Journal of Logic, Language and Information*, 13(2):207-223, 2004.

[7] Lance J Rips. The psychology of proof, 1994.

[8] Jakub Szymanik. *Quantiffiers and Cognition. Logical and Computational Perspectives. Studies in Linguistics and Philosophy*. Springer, forthcoming, 2016.