POLSKIE TOWARZYSTWO SEMIOTYCZNE

STUDIA SEMIOTYCZNE

Tom XXXIII • nr 1

PÓŁROCZNIK

MEANING, CONTENT AND REFERENCE

WARSZAWA • 2019

Adres redakcji: Krakowskie Przedmieście 3, 00-047 Warszawa
e-mail: studiasemiotyczne@pts.edu.pl
http://studiasemiotyczne.pts.edu.pl/

ISSN 0137-6608; e-ISSN 2544-073X

Prace redakcyjne związane z przygotowaniem składu elektronicznego, korektą językową i zamieszczaniem artykułów naukowych na stronie internetowej „Studiów Semiotycznych" i na stronie internetowej „Studiów Semiotycznych—English Supplement” – zadanie finansowane w ramach umowy 803/P-DUN/2018 ze środków Ministra Nauki i Szkolnictwa Wyższego przeznaczonych na działalność upowszechniającą naukę.

Ministerstwo Nauki
i Szkolnictwa Wyższego

CONTENT

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 5–8
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.01

From the Editors

TADEUSZ CIECIERSKI,* PAWEŁ GRABARCZYK **

FROM THE ISSUE EDITORS

In June 2018, the second *Context, Cognition and Communication* conference was held in Warsaw. Philosophers and linguists from over twenty countries presented more than eighty papers on the broadly understood philosophy of language and mind. After the conference, we announced the call for papers for the special issue of "Semiotic Studies" under the general title *Meaning, Content and Reference*. The call was addressed both to conference participants and the philosophical community in general. The current issue of "Semiotic Studies" contains the best articles submitted in response to that call.

The issue opens with the paper by Mirco Sambrotta *Transparent Contents and Trivial Inferences.* In his paper, the author defends the view that we may, at least to some extent and in some cases, reconcile externalism with the thesis of transparency of mental content if we assume the inferentialist views on attitude's contents. According to that view in cases where there is a name-component in the content and the component is associated with *a priori* accessible application conditions we can say that the appropriate content is transparent. The application conditions in such cases are trivial but ontologically ampliative, that is they enable us to arrive at conclusions that are ontologically committed to the existence of certain kinds of objects.

Maria Dolores Garcia-Arnaldos's paper, *Content and Meaning Constitutive Inferences*, concerns the challenge of circularity which often emerges when we try to provide justification of logic referring to the meaning of logical terms. The best known example of this is the attempt to establish the rule of MPP adhering to inferences of the user. Even if we follow Boghossian and assume that infer-

* University of Warsaw, Institute of Philosophy. E-mail: taci@uw.edu.pl. ORCID: 0000-0001-5403-9590.

** IT University of Copenhagen, Digital Design. University of Lodz, Institute of Philosophy. E-mail: pagrab@gmail.com. ORCID: 0000-0002-1268-7474.

ences can be accepted in a blind (pre-logical) way and that this acceptance is sufficient to establish MPP, the move from particular inferences to MPP is possible only thanks to MPP. Needless to say, this results in circularity. Another well-known problem of inferential theories (or the conceptual role semantics in general) stems from their holistic nature. If the meaning of an expression depends on the inferential structure of the beliefs of its user, then every user ends up having their own meaning (as it is fairly improbable for them to have the same sets of beliefs). As pointed out by the author, many authors (Boghossian included) have tried to solve this issue using a revamped version of the analytic/synthetic distinction. But deriving meaning from understanding of the rules is not the only route inferential theories can follow. As pointed out by Maria Dolores Garcia-Arnaldos, some of the authors prefer to make the starting point the rules themselves.

In *The Liar, Contextualism, and the Stalnakerian View of Context*, Jakub Rudnicki analyzes attempts to deal with (at least some versions of) the Liar Paradox by stressing the context-sensitivity of the liar sentence and the truth predicate. Rudnicki discusses briefly the proposal laid out by Michael Glanzberg who argues that we cannot reconcile the idea of representing contexts as context sets and basic observations about the liar without giving up on the idea of modifying a single set of possible worlds during the successive stages of conversation. Rudnicki questions this diagnosis and offers his own analysis based on the notion of s e m a n t i c d i s s o n a n c e which enables him to describe the reasoning leading to the Liar as enforcing the retraction of contextual update.

İskender Taşdelen tries to evaluate theories of meaning in the light of their compliance with Dewey's empirical requirements towards scientific theories. As pointed out by the author, most of the traditional theories of meaning struggle with this requirement as they tend to posit metaphysical entities, such as "meanings" or "ideas". One way out of this problem is to lean towards inferential theories of meaning (in the vein of Sellars, Brandom or Peregrin). Problem is, even though theories of this sort do not sin against empiricism, they are prone to the charge of circularity as it is difficult to explain inferences without resorting to the notion of meaning. The solution to this problem proposed by Taşdelen comes in two stages. In the first stage we have to disconnect the theory of meaning from the parallel theory of reference. This step is hardly controversial at this point as it can be seen as following on from the well-known considerations of Quine and Kripke. Where Taşdelen's proposition becomes very interesting is the second step in which the author proposes to treat meaning rules not as constitutive (or defining) rules of language but rather as auxiliary rules similar to strategic rules in games. The difference (as explained by Hintikka and Sandu) boils down to a difference between the rules employed by all players of the game (constitutive rules) and the rules of good or efficient players (strategic rules). The author proceeds with a detailed presentation of this alternative approach to meaning and shows how it can accommodate for phenomena which are traditionally difficult

to explain on the grounds of inferential theories - specifically the phenomenon of semantic change.

In *Untangling the Knot of Intentionality: Between Directedness, Reference, and Content*, Pierre Steiner addresses a problem of multiple characterization of intentionality in terms of aboutness, contentfullness and having the property of being representational. Since all three kinds of characteristics are common in the literature the question is do they concern a single notion of intentionality. The author's reply is negative. He proposes a hypothesis according to which we have to distinguish two concepts of intentionality: intentionality-T and intentionality-C. The former is intentionality qua object-directedness, the latter is intentionality qua contentfullness. Next, the author sketches the logical map of possible relations between the two concepts and proposes an interpretation of the main thesis of anti-representationalism as the claim that intentionality-C is neither sufficient not necessary for intentionality-T. In the last sections of the paper he presents a pragmatic theory according to which public language is constitutive for intentionality-C (which is necessary for intentionality-T) while the acquisition of the public language depends on intentionality-T as a property of "[…] deeds, non-linguistic acts and behaviour, historically and socially situated".

The next paper in the volume concerns the concept of metaphor. Richmod Kwesi starts the analysis with a critical discussion over Davidson's causal view on metaphor according to which metaphorical sentences have only literal contents while all the other "meanings" that metaphors are supposed to have can be explained in terms of causal effects of sentences on their interpreters. Kwesi argues that the Davidsonian causal account is inadequate as a theory that attempts to describe the systematic account of how figurative expressions work in communication. Kwesi contrasts three general models of figurativeness: the first (non-Davidsonian) attributes literal and metaphorical meanings to words, and claims that both are involved in the compositional generation of the literal and metaphorical meanings, the second (Davidsonian) denies the existence of the metaphorical meanings of words and the entire compositional generation of figurative meanings, the third – defended by the author – enables metaphorical meanings of compound expressions but denies that they are compositionally generated out of the literal meaning of non-compound expressions. The third (non-compositional) view shares literalism of the Davidsonian view when it comes to words and embraces the non-literalism of compound expressions. As such it comes out as the most promising view on metaphorical content when a general desiderata for each theory of metaphors are taken into account (Kwesi carefully describes such desiderata in the final section of his paper).

In *Lexical Concepts as Fluctuating Structures*, Alyona Budnikova presents a theory of lexical concepts conceived as semantic units conventionally associated with linguistic forms. The author argues that they can be identified with structures consisting of conceptual slots filled with various types of information about the referent with different structural weight. Next, a way of modelling the graded structure of lexical concepts is proposed by the author. Roughly speaking, the

model assesses the weight of each constituting structure of conceptual slots according to its relevance for defining purposes, frequency of contextual profiling and salience in derivation processes. The author presents the application of the model (to particular English nouns) as well as discusses briefly its merits and potential limitations.

In *Basic Concepts: A Cognitive Approach*, Wiesław Walentukiewicz attempts to critically analyze theories of conceptualization devised in psychology. He starts with a juxtaposition of the philosophical understanding of concepts and the way they have been defined in psychological literature. The author identifies that the psychological literature contains two main strategies for explaining the human ability to categorize objects. On one account, objects are categorized via perceived similarity. On the competing account, they are perceived on the basis of key properties associated with given categories. As pointed out by Walentukiewicz, both strategies lead to different descriptions of the details of the categorisation process. This is especially visible in how categories created by children are later developed into fully-fledged categories. In contrast to the prevalent approach found in psychology the author advocates the general similarity based approach. He claims that the logical properties of the relation of similarity make it better for the explanation of the existing empirical data.

We would like to thank the authors for their high quality articles, the reviewers for their insightful comments and finally the editorial staff of “Semiotic Studies” (Andrzej Biłat, Dominik Dziedzic) for their support during the work on this issue.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 9–28
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.02
Article

MIRCO SAMBROTTA*

TRANSPARENT CONTENTS AND TRIVAL INFERENCES

SUMMARY: A possible way out to Kripke's *Puzzle About Belief* could start from the rejection of the notion of epistemic transparency. Epistemic transparency seems, indeed, irremediably incompatible with an externalist conception of mental content. However, Brandom's inferentialism could be considered a version of externalism that allows, at least in some cases, to save the principle of transparency. Appealing to a normative account of the content of our beliefs, from the inferentialist's standpoint, it is possible to state that a content is transparent when name-components of that content are *a priori* associated with some application conditions and, at the same time, reflection alone provides an *a priori* access to those application conditions, with no need of any empirical investigation. Nevertheless, such requirements are only met in trivial cases. The aim of this paper is to argue that some application conditions of that sort, albeit trivial, can be ontologically ampliative. As a result, the related contents can be regarded as transparent in a substantial and rich way.

KEYWORDS: referentialism, transparency, externalism, application conditions, trivial inferences.

1. INTRODUCTION

Kripke's puzzle about belief shows the incompatibility between the principle of transparency and externalism. The principle of transparency states that anyone is in a position to notice and correct contradictory beliefs if one has them. If two of a thinker's token thoughts possess the same content, then the thinker must be able to know *a priori* that they do; and if two of a thinker's token thoughts pos-

* Universidad de Granada, Dpto. Filosofía I. E-mail: mirco.sambrotta@gmail.com. ORCID: 0000-0001-6962-4207.

sess distinct contents, then the thinker must be able to know *a priori* that they do. According to the principle of transparency, a subject cannot consistently believe a proposition and its negation at the same time, since their contradictory nature must be reflectively accessible to her with no need for any empirical investigation. The intuitive appeal of this principle stems from the intuitive appeal of the idea that anyone is in a position to know the contents of one's propositional attitudes. But certainly, this idea cannot be taken for granted in an externalist framework.

Content externalism claims that the contents of linguistic expressions are determined partly by certain (environmental or social) factors external to an individual speaker's inner state. Once we endorse this thesis concerning the contents of linguistic expressions, then it becomes almost unavoidable to endorse the same thesis concerning the contents of propositional attitudes, since those two sorts of contents are arguably dependent on each other.[1] The external factors that the thesis claims partly determine the contents of our statements and propositional attitudes can be completely unknown to us. Consequently, the thesis of content externalism entails that the contents of linguistic expressions and propositional attitudes are in a sense beyond our *a priori* grasp. Nevertheless, transparency of propositional attitudes seems to demand the opposite. If the contents of propositional attitudes are determined partly by certain external factors, of which we can be completely unaware, then the transparency principle is threatened. Therefore, epistemic transparency seems irremediably incompatible with an externalist conception of mental content.

It is in principle possible to give a normative account of the content of our beliefs. The inferential semantics can be considered as a paradigmatic example of such an account. According to the semantic inferentialist's standpoint, the content of a belief consists in certain inferential relations. The content of a belief, or the content expressed by a corresponding assertion, is given by the inferences the speaker is committed to and the justifications that entitle the speaker to make the assertion. Speaking of contents of beliefs is, therefore, speaking of commitments and entitlements. These norms of commitments and entitlements may also be defined in terms of exit-rules for name-components (noun terms) of the prop-

[1] In this paper I avoid complication due to context-dependence in natural language and I endorsed what Recanati (2018) refer to as the "Simple view", according to which the central meta-semantic question about the relation between the notion of content used in belief-desire psychology and the notion of content or meaning applied to expressions of the language has a simple answer. Roughly, by uttering a sentence that means that *p*, the speaker expresses her belief that *p*. So there is a single entity which is both the content of the belief (expressed by the speaker, and communicated to the hearer) and the content of the sentence. The entity which is both the content of a (declarative) sentence and the content of the corresponding belief is a proposition. What is important about propositions, however we analyze them, is that they are truth-bearers: they are true or false. Accordingly, beliefs and sentences are truth-evaluable because they have contents (propositions).

ositions believed (which tell us what we are committed and entitled to on the basis of applying a term). However, other meaning-constituting rules for name-components of the propositions believed may also be involved, including not only exit rules but also application conditions which serve as something like introduction or entry rules. Given such a normative standpoint, it is possible to maintain that a content is transparent when it is *a priori* associated with application conditions of that sort and then reflection alone provides an *a priori* access to those application conditions, with no need of any empirical investigation. This is an account of transparency that also externalists may accept, although such requirements are taken to be met just in trivial cases. Application conditions are taken to be trivial (in the sense of not requiring substantive investigation) when they reflect conceptual truths. Conceptual truths are articulations of constitutive semantic rules that govern proper use for the very noun terms we master as we acquire language. They are known *a priori* in the sense that competent speakers are licensed, based on their competence, to accept them, since they are just object-language expressions of rules they master. Rules of use entitle us to make trivial inferences, which can be considered as illustrations of such rules.

In what follows I shall try to argue that some application conditions of that sort, albeit trivial, are ontologically ampliative and, as a result, the related contents can be regarded as transparent in a substantial and rich way. They are existence entailing (ontologically ampliative), in a minimalist or "easy" approach to ontology, in the sense that beginning from an undisputed claim that makes no mention of a kind of entity *F*, we end with a claim that there are *F*s just by the undertaking of trivial inferences. Given an undisputed truth and by making use only of trivial reasoning, competent speakers are entailed to reach ontological conclusions that there are new entities not referred to in the undisputed truth. In this way, the existence of the entities in question can be known *a priori* insofar as the truth of the ontological claim may be inferred by any competent user of the term (who has mastered the relevant trivial inference) without the need for knowing any empirical truth (since one may begin the inference from conceptual truth).

In Sect. 2, I will present Kripke's puzzle about belief. In Sect. 3, I will argue that, in this puzzle, Kripke makes an assumption which conflicts with his referentialism: such an assumption is what Boghossian has called epistemic transparency. In Sect. 4, I will focus on the incompatibility between epistemic transparency and the externalist conception of mental contents. In Sect. 5, I will introduce the semantic inferentialist's account of the content of our beliefs. In Sect. 6, I will conclude by rejecting the broader view that nothing is epistemically available simply on the basis of linguistic and conceptual competence.

2. THE PUZZLE

In *A Puzzle About Belief* (1979, pp. 102–148) Kripke tries to disarm challenges to direct reference theories of proper names that are based on the apparent

failure of substitutivity in propositional attitude contexts. Kripke's puzzle is aimed at a variant of Frege's puzzle, which traditionally has been used to criticize referentialist views (Millianism). Kripke tries to draw the sting from Frege's puzzle by creating a similar paradox, but one which does not rest on Millianism in any way. The idea is, indeed, to show that it is illegitimate to blame the paradox on Millianism. Kripke's main contention is that the puzzle shows that the substitution-failures in propositional attitude contexts that are normally blamed on a substitutivity principle licensed by referentialism can be generated without using any such substitutivity principles. If correct, this contention would seem to disarm the argument from substitutivity failure as a criticism of referentialism.

In the original puzzle, Kripke constructs a situation in which the propositions expressed by the embedded sentences in belief ascriptions are contradictory. I won't present the details of Kripke's well-known article, which I assume to be familiar to the reader; I will just remind that Pierre in Kripke's example assents to two sentences that, unbeknownst to him, seem to contradict each other, namely that "London is not pretty" and "Londres est jolie". The possibility of this being an accurate belief ascription is then challenged by Kripke on the basis that Pierre would be sufficiently rational not to believe contradictory propositions.

The puzzle rests on two principles. The first is the disquotation principle (DP), which states that "if a normal English speaker, on reflection, sincerely assents to '*p*', then he believes that *p*" (p. 112). Kripke also states a biconditional form of the DP, namely that "a normal English speaker who is not reticent will be disposed to sincere reflective assent to '*p*' iff he believes that *p*" (p. 113). The biconditional DP implies that failure to assent indicates lack of belief, as assent indicates belief. The second principle that Kripke states is the principle of translation (TP). It states that "if a sentence of one language expresses a truth in that language, then any translation of it into any other language also expresses a truth (in that other language)" (p. 113).

It would seem that Pierre holds both beliefs, therefore, that he has contradictory beliefs. But, this option seems to lead to insuperable difficulties. We can assume that Pierre is a leading philosopher and logician, and "surely anyone, leading logician or no, is in principle in a position to notice and correct contradictory beliefs if he has them" (p. 122). In brief, Kripke's puzzle attempts to arrive at a contradiction by stipulating that a subject is rational and then showing how the DP and TP lead to the subject having contradictory beliefs. This is supposed to be irrational, and hence a paradox arises. Kripke also constructs the so-called Paderewski-puzzle. This is used to show that the above problem can also arise within a single language, using phonetically identical tokens of a single name.

According to referentialism, the sole semantic function of a proper name is simply to refer; in other words, the view that what a singular term *a* contributes to determining the proposition expressed by "*a* is *F*" is simply its referent. It follows that if referentialism is true, then so is substitutivity:

(SU) If "*Fa*" is a sentence containing a referring term *a*, then substituting *a* by a referring term *b* does not change the truth-value of "*Fa*", if *a* and *b* have the same reference.

While Kripke has pointed out the difficulties of an indirect theory of reference for names, he does not offer a solution to the problem of substitution of co-extensive names in belief contexts. Nevertheless, in order to links rationality with the absence of logically contradictory belief, we need to postulate another auxiliary principle, consistency:

(CO) If a speaker *S* reflectively and occurrently believes that *a* is *F* and that *a* is not *F*, then S is not fully rational.

The foregoing suggests that since (SU) follows straightforwardly from the referentialist semantics, because it is semantically irrelevant how that proposition is referred to, something better be amiss with (CO). The idea is then that if *S* is as rational as anyone gets, then *S* cannot hold contradictory beliefs. Reflection seems to be the operative principle behind consistency: since one can by reflection alone determine that one of one's occurrent beliefs is the negation of another of one's occurrent beliefs, if one is rational then upon reflection one should be able to detect the contradiction and thereby reject at least one of the beliefs.[2] No doubt most of us hold some contradictory beliefs without thereby being irrational, but we tend to think that had our cognitive abilities been as good as they get, we would not have held such belief.

Respectively, if the direct reference theorist does maintain this position, then we are left with the conflict between our logical instincts (if *a* equals *b*, then *b* can be substituted for *a* in a sentence without loss of truth) and our common sense (utterances may differ in truth value, because they may express different propositions). Nevertheless, I shall try to argue that the problem does not, strictly speaking, lie with the direct reference theory of names, but rather in the traditional view of believing.

3. TRANSPARENCY OF INCONSISTENCIES

The referentialist, however, may reject consistency: if the logical properties of belief content are not reflectively accessible to *S*, then *S* can hold contradictory beliefs without being irrational. The fact that Pierre's beliefs have logically contradictory properties is not reflectively accessible to him; it can only be discovered by appropriate empirical investigations. It means that Pierre lacks reflective access to key logical properties of the sole propositional contents of his

[2] Similar results hold for other logical relations among thoughts: e.g., if thought A is the negation of thought B, then I can know by reflection alone that thought A is the negation of thought B.

occurrent beliefs, that is, he is ignorant of such basic inferential relations between them as identity or contradiction. Therefore, *a priori* reflection will not be sufficient for him to amend his error; what he needs to discover is that he is thinking about two different individuals, which he can learn only by empirical investigation.

A number of philosophers have pointed out that Kripke makes an assumption which conflicts with his referentialism. This assumption is what Boghossian (1994, p. 33) has called "epistemic transparency". He formulates it as follows: Epistemic content is transparent if, and only if

> [when…] two of a thinker's token thoughts possess the same content, then the thinker must be able to know *a priori* that they do; and (b) if two of a thinker's token thoughts possess distinct contents, then the thinker must be able to know *a priori* that they do. (1994, p. 36)

If a person knows *a priori* whether or not the propositions expressed by two token thoughts are the same, then the logical properties of such propositions, i.e. whether they are consistent or contradictory, are similarly known by him *a priori*. Epistemic transparency would imply that determining and correcting contradictory beliefs is a matter of logical acumen, rather than acquiring information. But, as stated by Kallestrup: "[t]he fact that her beliefs have logically contradictory properties is not reflectively accessible to him; it can only be discovered by appropriate empirical means" (Kallestrup, 2003, p. 112).

On the contrary, epistemic opacity (the denial of epistemic transparency) would imply that logical acumen is not sufficient to detect all contradictory beliefs. The person would not be in a position to determine and correct all potentially contradictory beliefs unless he has acquired information regarding the content of the terms he used and thereby gained knowledge of the logical properties of the propositions expressed by two given sentences. What epistemic opacity denies is that propositional content is transparent in the sense that if *S* fully apprehends two propositional contents with a certain logical property, then *S* can come to know just by reflection that they have that property. Only then we accept epistemic opacity, *S* can consistently believe a proposition and its negation at the same time, since only then their contradictory nature is not reflectively accessible to her (Salmon, 1986, p. 132).

Referentialism and epistemic transparency turn out to be jointly inconsistent and what must be rejected, it seems, is epistemic transparency. This means that the referentialist should reject consistency, at least on our assumption that (CO) pertains to the propositions expressed by the embedded sentences, i.e., that if S occurrently and on reflection believes a proposition and its negation, then *S* is less than fully rational. Rejecting epistemic transparency would entail that also "a leading philosopher and logician" (Kripke, 1988, p. 123) can have contradictory beliefs, without this being irrational. Therefore, the most basic cleavage when considering the semantics of belief-attribution turns out to be between theories that claim content to be transparent and theories that do not.

Let me sum up. The paradox poses a difficulty for referentialism, because of its adherence to (SU). This reasoning has been argued to be fallacious based on the assumption of epistemic transparency. However, the referentialist may reject (CO), which relies on epistemic transparency, and so avoid inconsistency. If epistemic transparency is rejected, then it is a trivial matter to construct cases where *S*'s dissent from "*a* is *F*" does not imply that it is not the case that *S* believes that *a* is *F*. The implicit assumption of epistemic transparency is something that is peculiar to Kripkean puzzles; if epistemic transparency is refused, then the referentialist can easily avert the puzzle.

4. CONTENT EXTERNALISM

A referentialist view of the semantic content of proper names seems to represent a paradigm case of externalist content leading to epistemic opacity. Both Kripkean and Fregean puzzles involve a situation where an externalist would contradict the truth-value that the subject of a belief-attribution would assign to the belief-attribution. The referentialist can attribute contradictory beliefs to Pierre, safe in the knowledge that such attributions are made possible by the very nature of externalism. Epistemic opacity would seem to be a direct implication of holding an externalist conception of mental content. In such a conception of mental content, "[s]ubject's intentional states are individuated in part by certain sorts of facts about the physical and/or social environment in which he happens to be situated" (Boghossian, 1994, p. 34).

If my intentional states are individuated in terms of external facts (physical or social), then I cannot determine the logical properties of propositions expressed by token thoughts that are individuated in such a way without reference to, and knowledge of, these external facts. In other words, externalism would imply that determining the consistency of two token thoughts is sometimes a matter of acquiring information after all, and not only a matter of logical acumen. Boghossian, for instance, concludes that externalism entails a rejection of epistemic transparency, and takes this to be one of the main conclusions to be drawn from Kripke's puzzle:

> Now, it is fairly easy to show that externalist conceptions of mental content do not satisfy the transparency of sameness. Kripke's notorious Frenchman, Pierre, already shows this for the special case of Millian contents (themselves, of course, a sort of externalist content). (Boghossian, 1994, p. 36)

Therefore, why are these semantical facts puzzling? They are puzzling because of transparency of inconsistency: that anyone is in principle in a position to notice and correct contradictory beliefs if he has them. So logic alone should teach Pierre that his beliefs are inconsistent, yet it cannot. No logical reflection can show him the inconsistency. The intuitive appeal of this principle draws from the intuitive appeal of the idea that anyone is, in principle, in a position to know the contents of one's propositional attitudes. But certainly, in the framework of

externalist semantics, this idea cannot be taken for granted. Under the externalist supposition that the contribution of the name-components to the propositions believed is fully determined by the identity of the bearers of the names, then, those semantical data should be regarded as puzzling only once a compelling argument for the validity of this principle in an externalist framework is suggested.

Nevertheless, there may be convincing reasons why epistemic transparency is worth preserving. Usually, we think that it is an essential ingredient in what has come to be known as "privileged access": the idea that *S* has a first-person authority with respect to the contents of her own occurrent mental states. If the subject lacks reflective access to key logical properties of the sole propositional contents of her occurrent beliefs, that is, she is ignorant of such basic inferential relations between them as identity or contradiction. This, in conjunction with the fact that our ordinary way of talking when attributing beliefs always agrees with the truth-value assignment made by any rational subject of a belief-attribution, then suffices to generate problematic belief attributions. In fact, the falsehood of the intuitive reflection principle, that one can by reflection alone determine the simple logical relations among one's propositional attitudes, seems to conflict with our ordinary intuitions regarding belief. People commonly and without hesitation do accept the truth of the datum. This means that our common practices of belief-attribution treat the content embedded in propositional attitude contexts as epistemically transparent.[3] Therefore, externalist theories will result in attributions that contradict our normal practice of belief-attribution and, accordingly, we can characterize the problem cases as those where an externalist would contradict the assignment of truth-values of a rational agent.[4]

We can differ from the truth-value assignments of a rational agent if the content is externalist, while internalist theories need not have these results. Internalism of this class thus decrees that some mental states are transparent in the general sense that they are fully "open to view" or "revealed by introspection". From an internalist standpoint, it is natural to maintain that in some instances there is

[3] However, our common practices of belief-attribution, like accepting the datum, only show that our common practices are committed to epistemic transparency, not that it is correct. This raises an intriguing possibility: what if our folk semantics is internalist, descriptivist and epistemically transparent, but actually wrong (and/or incoherent) in some sense?

[4] Against the intuition that the job of semantics simply is to treat our common attribution of truth-values as data, various referentialists, for instance Salmon (1986), have argued that the datum is, in fact, false and that our ordinary way of speaking is to be explained with reference to Gricean implicatures and the like. Salmon claims that we often take a true sentence like "Lois Lane believes that Clark Kent can fly", which actually expresses a true singular proposition, to include the Gricean implicature that Lois Lane would assent to "Clark Kent can fly". Here the implicature of the sentence would "lead speakers to deny it, despite its literal truth-conditions [being] fulfilled" (Salmon, p. 115).

a necessary connection between believing a proposition and its truth.[5] This distinction and not the descriptivist/referentialist distinction could be considered the fundamental distinction in the semantics of belief-attributions.

In the same line, David Lewis states:

> When we characterise the content of belief by assigning propositional (or other) objects, are we characterising an inner, narrowly psychological state of the believer? Are beliefs in the head? Or are we characterising partly the believer's inner state, partly the relations of that state to the outer world? If it is the latter, the objection may succeed; however, Kripke's puzzle vanishes. For if the assignment of propositional objects characterises more than the believer's inner state, then there is no reason to suppose that a leading philosopher and logician would never let contradictory beliefs pass, or that anyone is in principle in a position to notice and correct contradictory beliefs if he has them. Anyone is in principle in a position to notice and correct a state of the head which can be characterised by assigning contradictory propositional objects, but why should philosophical and logical acumen help him if the trouble lies partly outside? (Lewis, 1981, pp. 288–289)

Briefly, if we agree with this kind of externalism that *a priori* reflection will not always suffice to ensure the validity of our inferences, then it looks like that embracing externalism and abandoning transparency "blurs the distinction between errors of reasoning and errors of fact" (Boghossian, 2011, p. 458). The point is that, if it were true that *a priori* reasoning did not suffice for avoiding logical errors, then we could not assure our status as rational agents by mere *a priori* reflection (and this last point is an important thesis for how we have traditionally understood what it is to be rational). But, even though we accept that content is not transparent and, therefore, that one might be condemned to make logical mistakes, we want, and must, consider Pierre to be a rational person!

5. CONTENT OF BELIEF

What is the content of a belief expressed by a sentence with a proper name? It is very difficult to have a uniform conception of the content of a belief and it is exactly the main problem arising from the discussion of Kripke's puzzle. A relevant assumption, which does not enter directly into the argument of the puzzle, works as the background for Kripke's assumptions about the propositional content of a belief:

[5] However, internalism is a highly contentious position, reproved by the likes of Burge, Putnam, Kripke, and Williamson.

- Proper names are a basic ingredient in forming singular propositions, intended in a Russellian way; therefore the content of belief, or of other propositional attitudes, is given by an ordered pair with an object and a property; in the relevant cases—London, ugliness/prettiness—or—Paderewski, musical talent.

At the end of his paper, Kripke speaks of the "cloud" that the paradox places over the notion of content and this cloud seems to be placed mainly on the idea of a singular proposition. In general, if "*a* is *F*" expresses a singular proposition, then "*S* believes that *a* is *F*" is true iff "*S* believes of *a* that it is *F*" is true. According to millianism, singular propositions are individuated by their objectual constituents, independently of how they are conceptualized. An utterance of "*a* is *F*" thus expresses a singular proposition consisting of the referent of *a*, and an utterance of "*b* is *F*" expresses the same proposition if $a = b$. Likewise, the sole propositional content of a belief is a singular proposition consisting of an object and a property which are what the embedded sentence "*a* is *F*" refers to and it is semantically irrelevant how that proposition is referred to**.** Hence, contradictory beliefs are between singular propositions. It is also possible to assume a relational principle of belief on the referentialist account, according to which S believes that "*a* is *F*" says that *S* is belief-related to a singular proposition B(*S*; (*a*, *F*-ness)) and the only propositional content of *S*'s belief is a singular proposition. It is therefore not to be expected that speakers who entertain such propositions can come to know their logical properties "just by deploying their conceptual apparatus from the armchair" (Kallestrup, 2003, pp. 112–113). If so, then the content of one's belief would still be unknown to the subject, and transparency would not be preserved, after all.

Nevertheless, the answer to this question can be given adequately also in the spirit of the use theory of meaning, from the semantic inferentialist's standpoint. Roughly, the content of what we say and judge is inferentially articulated by being pragmatically determined in normative practices of scorekeeping. In accordance with Brandom's inferential role semantics, the content of a belief, or the content expressed by a corresponding assertion, is given by the inferences the speaker is committed to and the justifications that entitle the speaker to make the assertion. Speaking of the content of a belief is, therefore, speaking of commitments and entitlements. Propositional contents consist in their distinctive role in inferences and can be identified with the inferential relations one is committed to, or with the inferential commitments one undertakes in expressing a claim or a belief.

From the inferentialist semantic standpoint, the representational aspect of the propositional content of a claim consists of the inferential roles of various true identity statements that describe the identity condition of the object. Inferential

roles are enabling us to make new substitutional commitments[6] through substitutional inferences, i.e. the inferences that draw a consequence by simultaneously replacing a certain term occurring in a premise with another term based on an identity statement. Therefore, grasping the content of a statement consists of the ability to derive various substitutional commitments from the original statement together with those true identity statements making through substitutional inferences. By analogy, contents of beliefs can be also explained in a similar way.[7] The occurring terms in belief ascription, as interpreted by the ascriber, reflect the ascriber's own acknowledged substitutional commitments, hence what the ascriber takes to be the objective substitutional norms governing these terms. And since objective norms bind everybody, the ascription thus captures substitutional commitments that, according to the ascriber, the ascribee is bound to acknowledge given the belief ascribed, whether or not the ascribee acknowledges them in fact.

One of the main tenets of *Making it Explicit* (Brandom, 1994) is that the constituents of propositional contents are the objective semantic norms governing the use of among speakers of English: this commits Brandom to a version of semantic externalism (Brandom, 1994, p. 632; 2000, pp. 359–360). In this sense, the content of a statement or belief is not only inferentially but also socially articulated in our inferential practice. In order to grasp the representational aspect of the propositional content of a statement or a belief, we should attend not merely to the inferentially articulated dimension, but also to the socially articulated dimension of our game of giving and asking for reasons. Moreover, a common content has to be considered not as something shared by every member of the society, but as generally accepted norms towards which all people should conform and do conform when properly guided.

In particular, according to inferentialism, propositions form equivalence classes under substitution in good material inferences. Once this is acknowledged and inferential roles are defined, it's easy to realize that the same substitutional strategy can be applied to obtain indirect sub-propositional inferential roles. Brandom's idea is precisely that keeping the score of simple material substitutional inferential commitments (SMSICs) requires treating speakers as committed to certain tokenings being part of certain recurrence structures which behave as anaphoric chains: a certain tokening is governed by certain SMSICs to the extent that it treated as part of a certain anaphoric chain. Brandom defines a n a p h o r i c c o m m i t m e n t, a commitment to treat tokening as belonging to a certain anaphoric chain. Surely anaphoric chains may extend not only intrapersonally but also interpersonally among tokenings by different speakers. This is crucial for the objectivity and the s c r u t a b i l i t y of reference. In fact, to treat one's own tokening as anaphorically dependent on another one's tokening

[6]A substitutional commitment of a claim is the commitment undertaken by an interlocutor from *de re* viewpoint (see Brandom, 1994, pp. 136–140, 370–376, 495–520).

[7] The content of a belief can be defined as the commitment to correct substitutions with respect to anaphoric chains.

is to treat it as governed by the same SMSICs that govern the antecedent. Of course, speakers might get anaphoric commitments wrong, as it happens in the case of Kripke's puzzle.

Remind that Pierre in Kripke's example assents to two sentences that, unbeknownst to him, seem to contradict each other, namely that "London is not pretty" and "Londres est jolie". The possibility of this being an accurate belief ascription is then challenged by Kripke on the basis that Pierre would be sufficiently rational not to believe contradictory propositions. In Brandom's framework, the contradiction is to be explained away simply pointing out that Pierre in fact does not realize that London_1 and Londres_2 (or Paderewski_1 and Paderewski_2) belong to the same anaphoric chain. This solution of the puzzle, of course, questions Kripke's tenant that when one finds oneself holding a contradiction as a consequence of the application of the principles of translation and disquotation, "logic alone should teach him that one of the beliefs is false" (Kripke, 1979, p. 399). In line with such a position, Pierre's two beliefs differ in their inferential roles. The substitutional commitment is undertaken through the substitutional inference based on an identity statement to which the interlocutor is committed, regardless that the speaker may acknowledge the commitment or not. Accordingly, Pierre erroneously attaches to the same type of name two different sets of inferential commitments, which are two different ways to keep track of the one individual, without acknowledging it. Pierre has, therefore, two different contents of belief, even if his beliefs are actually connected to the same referent (to the same object of belief).

6. Application Conditions and Trivial Inferences

As we have seen, roughly speaking the content of a propositional attitude is transparent if there is no significant gap between the thought and what the thought is about. Unfortunately, the transparency seems incompatible with externalism conception of mental contents. The externalist cannot possibly maintain that contents are transparent due to some key logical properties, such as identity and contradiction, which are not immediately revealed by them. According to such a conception, contents are related to reality by facts external to our *a priori* grasp, hence it do lead us to deny that there can be any transparent contents. But, our terms also come with rules of use we master as we acquire language. Along with inferentialist account, we can speak more generally of the introduction rules for terms. In some cases, the introduction rules may license us to apply a term if certain application conditions[8] are fulfilled. If one takes this general approach, application conditions for nouns can be treated as among the introduction rules licensing us to apply certain terms. All that is required here is that a content is

[8] Application conditions can be, for instance, conditionals of the form "If *x* is *P*, then *x* falls under concept *C*" (or "If *x* falls under *C*, then *x* is *P*"), where *P* is some property (or set of properties) of the object *x* (see Diaz-Leon, 2014, p. 12; Levine, 2001).

transparent when it is *a priori* associated with application conditions of that sort and then reflection alone provides access to the application conditions for that content with no need of any empirical investigation or to be supported by empirical evidence about the relevant external conditions.

This is an account of transparency that also externalists may accept, although such requirements are taken to be met just in trivial cases. That may be right, but it does not mean they never truly count as revealing anything. Indeed, some application conditions of that sort, albeit trivial, can be ontologically ampliative, and hence the related contents can be regarded as transparent in a substantial and rich way.

Application conditions are taken to be trivial (in the sense of not requiring substantive investigation) when they reflect conceptual truths. Conceptual truths are articulations of rules of use for the introduced noun term. Application conditions of that sort are among the constitutive semantic rules that govern proper use for the very noun terms we master as we acquire language. Rules of use entitle us to make trivial inferences, which can be considered as illustrations of such rules. Therefore, conceptual truths may be seen as object-language articulations of the rules that may be used in introducing terms.

A trivial inference that relies on application conditions of that sort can be ontologically ampliative without being informationally ampliative.[9] They are existence entailing (ontologically ampliative), conforming to a minimalist[10] or "easy" approach to ontology, in the sense that beginning from an undisputed claim that makes no mention of an entity *F*, we end with a claim that there are *F*s. That is, we obtain a derived claim which entails the existence of *F*s, just by undertaking and making use of trivial inferences. In other words, given an undisputed truth and by engaging in trivial inferences, we can reach a truth that is intuitively redundant with respect to the first one, which yet leaves us with ontological commitments to disputed entities (Thomasson, 2015, p. 234). The point is that we can use trivial inferences of that sort to acquire commitments to trees, tables, volcanoes or any ordinary object if we start (in a metaphysical dispute, for instance) from an undisputed claim such as "there are particles arranged volcano-wise". For it is a conceptual truth (a truth knowable *a priori* via command of the

[9] David Chalmers speaks of an inference as ontologically ampliative if, roughly, "the consequent makes an existential claim that is not built into the antecedent" (Chalmers, 2009, p. 95). It is worth noting that this use of "ampliative" is crucially different from Contessa's sense of informationally ampliative: the inference from (a) to (b) is informationally ampliative: when (b) contains new empirical information not present in (a) (Contessa, 2016, *passim*; see also Thomasson, 2017, p. 771)

[10] The expression "ontological minimalism" is taken from the work of Thomasson (2001), where she uses it to describe her own and Schiffer's view. Since Thomasson has moved away from that name and prefers "easy ontology". Her reason for moving away from ontological minimalism is that it suggests, not her view that the standards for ontological commitment are minimal, but that the entities that exist, according to the view, are somehow minimal, an implication she rejects.

term) that the existence of volcanoes is guaranteed whenever there are particles arranged volcano-wise (Schiffer, 2003, p. 52; Thomasson, 2015, p. 149, 231).

In Schiffer's terms, they are "pleonastic something from nothing inferences". Engaging in pleonastic something from nothing inferences, we begin with undisputed truths and combine it with an analytic or conceptual truth that functions as what Schiffer calls a "transformation rule", to give us a derived claim that is, intuitively, redundant with respect to the undisputed claim, yet leaves us with (apparently new) ontological commitments to the disputed entities (Schiffer, 2003, p. 52). Versions of inferentialism make use of such a kind of trivial inferences, endorsed by Schiffer and Thomasson, in developing the easy approach to ontology. According to easy ontological views, many ontological debates may be resolved by engaging in inferences that seem redundant in ordinary English as genuinely trivial from uncontroversial premises. The view is motivated by its ability to tackle directly the question of how propositional thoughts about such objects are possible and how they can be knowledgeable. For given the trivial inferences that take us to claims about objects, we can see how speakers may acquire knowledge of these objects by knowing the uncontroversial truths and mastering the rules of use for the terms that entitle them to make inferences from those uncontroversial truths to the existence of them (Thomasson, 2015, chap. 3, pp. 127–160).

What the easy ontologist needs is clearly a normative claim, about what competent speakers are entitled to conclude (and what would be a mistake), not a descriptive one about what competent speakers will be disposed to assent to. That is to say, a normativist version of inferentialism which treats possessing a concept not as entailing that speakers are disposed to assent to certain statements, but rather that they ought to assent. Inferential rules (typically expressed by conditionals, material or formal) do not primarily consist in obligations for speakers or believers; they rather constraint our linguistic practices by delimitating what, from an inferential point of view, we may and may not do by entertaining propositional contents. They should better be seen as normative uniformities characterizing the pattern-governed behaviours of speakers. The view is not that someone's understanding the claim entails that she has a disposition to assent to it, but rather that mastery of the relevant linguistic rules governing the expressions used entitles one to make the relevant inference using those expressions and embrace the ontological conclusion and that rejecting it would leave one open to rebuke.[11]

[11] There seem to be two separate questions here: how can an individual be obliged to reason according to certain rules, and why ought we (collectively) to have those rules rather than any others. On the first, the right approach seems to be that one can be so obliged by presenting oneself as a participant in the relevant public norm-governed practice (just as one can be obliged to follow the rules of soccer by joining the soccer game). The question of why we ought to adopt certain rules (or norms) rather than any others is far more difficult. One might look to the work of inferentialist logicians (Beall & Restall, 2013; Ripley, 2013) for a way of understanding certain basic norms regarding acceptance

In cases in which scientific investigation *seems* to have disproven common sense claims, metaphysicians have often offered eliminativist paraphrases. Eliminativists attempt to avoid commitment to disputed entities by translating all the ordinary language claims apparently requiring quantification back into statements that (try to) avoid such quantification. Following the well-known Van Inwagen's example (1990), "There is a chair here" would be paraphrased as "There are particles arranged chair-wise here" thereby quantifying only over particles, not chairs, thus supposedly avoiding commitment to chairs. Nevertheless, the two sentences involve sortals terms which, although different, are found to have the same application conditions—where "application conditions" (Thomasson, 2015, p. 90) can be said to be among the semantic rules of use for the sortal terms we master as we acquire language that determine in which situations they are successfully applied. This is so because the sortal term for the given kind of entity "chair" may be derived simply by pleonastic transformations (Schiffer, 2003) from the basic sentences "There are particles arranged chair-wise here", which does not quantify over anything of that kind. In particular, from the language of refuge, the statement "There are particles arranged chair-wise" one could still form the nominalization "A chair-wise arrangement (of particles)" or, for short, "A chair." Accordingly, if the latter claim "There are particles arranged chair-wise here" turns out to be true, then so it will be the former claim "There is a chair here." For the fulfillment of the application conditions of the first sortal may be sufficient to the fulfillment of the application conditions of the other. In this case, we are genuinely introducing reference to a new entity, not just relabeling an old entity of the same sort and we could state the application conditions for "Chair" without appealing the existence of a chair or indeed of any object at all. These conditions may be stated in the form of rules that enable us to move from talk that did not make use of the relevant noun term (or any synonymous or co-referring term) to talk that does—though again, it's worth emphasizing that there is no requirement that these application conditions be stachair at all.

Therefore, the question "Is there a chair there?", can be straightforwardly answered by beginning from a claim that is not a point of controversy between realists and (most) eliminativists:

- Uncontroversial claim: "There are particles arranged chair-wise".

But, the following seems to be a conceptual truth:

- Conceptual truth: "If there are particles arranged chair-wise, then there is a chair".

---

and rejection as constitutive norms for thought, and thus as non-optional. And we might look for a pragmatic justification for adopting other (less basic) norms or rules.

Indeed, the occurrence of the situation in which eliminativists would say particles are arranged chair-wise guarantees that the application conditions for the ordinary term "chair" are met. Thus, competent speakers who master the application conditions for "chair" are licensed to infer the derived ontological claim:

- Derived/ontological claim: "There is a chair".

In this way, ontological debates about the existence of concrete objects can be settled just as "easily". Ontological claims may be derived by competent speakers, through inferences, from uncontroversial claims combined with conceptual truths. Accordingly, by trivial reasoning a speaker may entitle her to reach new conclusions. For mastery of the rules of use for terms license the speaker to make easy inferences from basic, uncontroversial truths to the existence of the entities in question and to move from knowledge of the conceptual truth to knowledge that the things in question exist.[12]

This seems perfectly consistent with our ordinary talk about existence (i.e., outside the ontology room): from the fact that there are (according to the eliminativist's theory) simples arranged chair-wise, we may infer that there is a chair. For according to ordinary usage, nothing more is required. Considered as part of normal English speech, the two sentences are inferentially bound to each other and the truth of the former is analytically sufficient (in the epistemic sense of being available to the speaker simply on the basis of linguistic/conceptual competence; Thomasson, 2007, p. 165) to ensure the truth of the latter—that is, epistemically available to the speaker simply on the basis of conceptual and linguistic competences. In this case, accepting the existence of a chair is analytically entailed by accepting the truth of the sentences describing the former state, therefore, accepting the truth of "There are simples arranged chair-wise" but denying the truth of "There are chairs" would be pointless. By treating the paraphrases as true and the direct claims about chairs as untrue, in fact, the eliminativist would sever trivial connections allowed in ordinary speech between sentences. As a consequence, treating the statements as lacking the same truth-value could only be done "by artificially inflating the application conditions for "chair" beyond those enshrined in normal use of the term" (Thomasson, 2007, p. 165).

To conclude, let's focus again on attitude reports. Imagine that Pierre finally arrives in London and enters a restaurant. Pierre has meanwhile become a mereological nihilist, so that he assents to the two following sentences: "There are particles arranged chair-wise" but "There is no chair". Accordingly, we can report:

i. Pierre believes there are particles arranged chair-wise.

[12] By adopting, perhaps, a minimalist approach to truth (see Price, 2011, pp. 253–279).

ii. Pierre believes that there is no chair.

However, given the aforementioned conceptual truth "if there are particles arranged chair-wise, then there is a chair", which is supposed to establish the application conditions of the noun term "chair" and to govern its use, if Pierre asserts that there are particles arranged chair-wise and we take him to be a competent speaker, then we are licensed to infer "Pierre believes that there is a chair". As a result, we end up attributing to Pierre both the belief in the proposition that "there is a chair" and the belief in its negation at the same time.[13]

But, as I tried to show, Pierre can consistently believe both propositions at the same time, only if their contradictory nature is not reflectively accessible to him. That is not going to happen here. In this case, unlike Kripke's original example, we can accept epistemic transparency. Indeed, the fact that Pierre's beliefs have logically contradictory properties should be accessible to him simply by reflecting on the way the word "chair" is actually used in linguistic practice and not through any empirical investigation. Mastering the rules of use of the noun term "chair" in linguistic practice is sufficient to determine that one of his occurrent beliefs is the negation of another of his occurrent beliefs. It means that *a priori* reflection would be sufficient for him to detect that he is holding contradictory beliefs and to amend his error. Therefore, if in this case the logical properties of the belief content are reflectively accessible to Pierre, then Pierre cannot hold such contradictory beliefs without being regarded as irrational (or, at least, as a non-competent speaker). That is, if Pierre occurrently and on reflection believes the proposition "There is no chair" and its negation at the same time, then we cannot take him to be fully rational (or rather, we cannot take him to be a competent speaker). Because if Pierre was rational (or better, a competent

[13] In order to make this case fit with referentialism and *direct reference* theories of proper names, we can imagine Pierre naming the particles arranged chair-wise in front of him as "Sum" and the alleged non-existent chair "Tab". Then we can report:

Pierre believes that Sum exists.

Pierre believes that Tab does not exist.

But, since the application conditions of Tab allows to assert that if there is Sum, then there is Tab, we are licensed to infer "Pierre believes that Tab exists" (or better, he should if he was a competent speaker). So, we end up by acknowledging that:

Pierre believes that Tab exists.

Pierre believes that Tab does not exist.

As a result, we end up attributing to Pierre both the belief in the proposition that "There is a chair" and the belief in its negation at the same time. But Pierre can consistently believe both propositions at the same time, only if their contradictory nature is not reflectively accessible to him. That is not going to happen here. In this case, unlike Kripke's original example, we can accept epistemic transparency. Indeed, the fact that Pierre's beliefs have logically contradictory properties should be accessible to him simply by reflecting on the way the word "chair" is actually used in linguistic practice and not through any empirical investigation.

speaker), then he should be able to detect the contradiction and thereby reject at least one of the beliefs.

In the light of the above, it is possible to conclude that sometime competent speakers are *a priori* licensed to accept conceptual truths (object-language expression of rules they master) and to underwrite trivial inferences which are ontologically ampliative (without being informationally ampliative).[14] Therefore, in those circumstances, they can reach and acquire ontological conclusions without the need for knowing any empirical truths.[15]

7. CONCLUSION

Kripke's Puzzle about Belief shows how externalism and transparency of contents are always mutually incompatible. As stated above, although the inferentialist standpoint can be understood as a sort of externalism, it allows some inferential relations be epistemically transparent (epistemically available to the speaker simply on the basis of linguistic and conceptual competence).[16] Such inferential relations, albeit trivial, turn out to be somehow existence-entailing (ontologically ampliative in Chalmers's terms [2009, p. 95]), that is, they can reveal ontological commitments—insofar as the conclusions seems to commit us to the existence of things of a *sort* not mentioned in the premises. So, even with-

[14] Once again, an Inference can be considered not informationally ampliative roughly if the conclusion does not add any empirical information to the premise.

[15] An example of a trivial inference that is not ontologically ampliative is the following: I accept that there are female foxes (uncontroversial truth), female foxes are vixens, then there are vixens. Even though this inference can be deemed as transparent, it is not existence entailing: female foxes just are vixens—we are not inferring the existence of anything new. The particular conceptual truths used in easy ontological arguments do raise distinct cause for concern**.** For these have a peculiar feature that inferences like the above do not: the easy ontologist's inferences are existence entailing in the sense that we begin from an undisputed claim that makes no mention of Fs (or any coextensive concept) and end with a claim that there are Fs, a new kind of entity not previously mentioned, or an entity of a different sort. This is crucial to the idea that some existence questions may be answered easily by making use of application conditions, without this turning out to be circular. Thanks to an anonymous referee for asking me to clarify this distinction.

[16] Note that the view is not that the speaker must necessarily be aware of the relevant linguistic/conceptual norms in the sense of being able to formulate them in object language, but rather that mastery of the relevant linguistic/conceptual rules governing the expressions used entitles one to make the relevant inference using those expressions (and embrace the ontological conclusions). If we focus on the linking conceptual truth rather than the inference, we can express this as saying that mastery of the relevant linguistic/conceptual rules entitles one to accept the conceptual truth (without the need for any further investigation), and that rejecting it would be a mistake. What we need is a normative claim, about what competent speakers are entitled to conclude (and what would be a mistake). That is to say, speakers who master the relevant conceptual/linguistic rules are entitled to make the relevant inference, and to accept the conceptual truth (and are open to reproach if they refuse to).

out endorsing internalism, we can conclude that there are some transparent contents. At the same time, this leads us to reject the broader externalist view that nothing ampliative is epistemically available to the speaker simply on the basis of linguistic and conceptual competence.

I started this paper by summarizing Kripke's puzzle about belief. I suggested that one of the main achievements of this puzzle is to show the incompatibility between Boghossian's principle of epistemic transparency and externalism of mental contents. According to the latter, contents are related to reality by facts external to our *a priori* grasp (epistemic opacity). As a consequence, there cannot be any transparent content.

I then focused on the semantic inferentialist's account of the content of our beliefs. I put forward the idea that, on the basis of this normative standpoint, a content can be taken to be transparent when name-components of that content are *a priori* associated with some application conditions and then reflection alone provides an *a priori* access to those application conditions, with no need of any empirical investigation. In this way, it was possible to provide an account of transparency compatible with externalism. An account, though, that works just in trivial cases. Trivial application conditions reflect conceptual truths. I argued that inferences that rely on application conditions of that sort, albeit trivial, can be existence entailing—according to a minimalist or "easy" approach to ontology. Accordingly, I concluded that some transparent contents (contents epistemically available simply on the basis of linguistic and conceptual competence) turn out to be, to some extent, ampliative.

REFERENCES

Beall, Jc., Restall, G. (2013). Logical Consequence. *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/entries/logical-consequence/

Boghossian, P. A. (1994). The Transparency of Mental Content. *Philosophical Perspectives*, *8*, 33–50.

Boghossian, P. A. (2011). The Transparency of Mental Content Revisited. *Philosophical Studies*, *155*(3), 457–465.

Brandom, R. (2008). *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press.

Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brandom, R. (1994). *Making It Explicit. Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Chalmers, D. (2009). Ontological Anti-Realism. In: D. Chalmers, R. Wasserman, D. Manley, (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (pp. 77–129). Oxford: Oxford University Press.

Contessa, G. (2016). It Ain't Easy: Fictionalism, Deflationism, and Easy Arguments in Ontology. *Mind*, *125*(499), 763–73.

Diaz-Leon, E. (2014). Do *a Posteriori* Physicalists Get Our Phenomenal Concepts Wrong? *Ratio*, *27*(1), 1–16.

Evans, G. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.

Field, H. (2000). A Priority as an Evaluative Notion. In: P. A. Boghossian, C. Peacocke (Eds.), *New Essays on the a Priori* (pp. 117–149). Oxford: Oxford University Press.

Harrison, B. (1982). Description and Identification. *Mind*, *91*(363), 321–338.

Kallestrup, J. (2003). Paradoxes About Belief. *Australasian Journal of Philosophy*, *81*(1), 107–117.

Kripke, S. (1979). A Puzzle About Belief. In: N. Salmon, S. Soames (Eds.), *Propositions and Attitudes* (pp. 102–148). New York: Oxford University Press.

Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.

Lewis, D. K. (1974). Radical Interpretation. In: D. Lewis (Ed.), *Philosophical Papers* (Vol. I, pp. 108–121). New York: Oxford University Press.

Lewis, D. K. (1981). What Puzzling Pierre Does Not Believe. In D. Lewis, *Papers in Metaphysics and Epistemology* (pp. 416–417). Cambridge: Cambridge University Press.

Price, H. (2011). *Naturalism without Mirrors.* Oxford: Oxford University Press.

Recanati, F. (2018). From Meaning to Content: Issues in Meta-Semantics. In: D. Ball, B. Rabern (Eds.), *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics* (pp. 113—137). Oxford: Oxford University Press.

Ripley, D. (2013). Paradox and Failures of Cut. *Australasian Journal of Philosophy*, *91*(1), 139–64.

Russell, B. (1912). *The Problem of Philosophy*. Oxford: Oxford University Press.

Salmon, N. (1986). *Frege's Puzzle*. Cambridge: MIT Press.

Schiffer, S. (2003). *The Things We Mean.* Oxford: Oxford University Press.

Thomasson, A. L. (2017). Why We Should Still Take It Easy. *Mind*, *126*(503), 769–779.

Thomasson, A. L. (2015). *Ontology Made Easy*. Oxford: Oxford University Press.

Thomasson, A. L. (2001). Ontological Minimalism. *American Philosophical Quarterly*, *38*(4), 319–31.

van Inwagen, P. (1990). *Material Beings*. Ithaca: Cornell University Press.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 29–47
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.03
Article

MARIA DOLORES GARCIA-ARNALDOS *

CONTENT AND MEANING CONSTITUTIVE INFERENCES[1]

SUMMARY: *A priori* theories of justification of logic based on meaning often lead to trouble, in particular to issues concerning circularity. First, I present Boghossian's *a priori* view. Boghossian maintains the rule-circular justifications from a conceptual role semantics. However, rule-circular justifications are problematic. Recently, Boghossian (Boghossian, 2015) has claimed that rules should be thought of as contents and contents as abstract objects. In this paper, I discuss Boghossian's view. My argumentation consists of three main parts. First, I analyse several arguments to show that in fact, Boghossian's inferentialist solution is not fully satisfying. Second, I discuss the matter further, if one accepts that basic logical rules are constitutive of meaning, that is, they constitute the logical concepts and the content of a rule is an abstract object, then abstract objects—like, for example, rules—could be constitutive of meaning. The question is whether conceptual priority is in the judgment or in the object and what theory of content is pursued. Grasping content as a matter of knowing how a word or concept behaves in inferences is not completely explicative. Finally, I contend that rules come to exist as a result of certain kinds of mental action. These actions function as constitutive norms. Logical rules are not abstract objects but ideal. What one construes as norms or rules of content may involve idealization, but this is because we share a language.

KEYWORDS: meaning, conceptual role semantics, inferentialism, content, logical rules.

* CEU San Pablo University, Faculty of Humanities and Communication Sciences. E-mail: dolores.garcia.arnaldos@usc.es. ORCID: 0000-0001-9998-8060.

[1] I appreciate all the help of the two anonymous referees whose criticism helped greatly in improving the arguments of this paper. Thanks, as well for their insightful indications and their linguistic suggestions.

INTRODUCTION

The main goal of this study is concerned with the meaning and the content of the basic logical rules. In particular, this work investigates—in the overall view of the contemporary theories in philosophy of logic—the idea of the normative content of a rule.

In recent years, philosophical reflection about the justification of logic has been dominated by the enduring problem of circularity and the infinite regress puzzle. *A priori* theories of justification of logic, based on meaning, are standardly recognized to be akin to the conceptual role semantic theories that often lead to similar problems; in particular, to issues concerning a vicious circle. In general, to justify the logical rules inferentially involves using logical principles, which generates circularity. How can we justify logical laws in a way that doesn't rely on those laws?

In his defence of an inferential justification, Paul Boghossian (Boghossian, 2000; 2001; 2003a; 2003b)[2] has argued that the only way to justify the basic logical principles is to claim a rule-circular justification, according to which, knowledge of the validity of the basic rules of inference is the result or product of inferences. Boghossian considers the *apriority* of logical propositions from a meaning-based approach, that is, from theories that defend the claim that the understanding of the meaning contributes to the explanation of the justification of beliefs or the transition between beliefs of a person. In this sense, he assumes that the principles of logic, as theoretical principles, are part of the meaning of language expressions. Logical constants[3] are implicitly[4] (tacitly) defined, taking logical basic rules into account, which allow their introduction and/or elimination. According to his proposal, some patterns of basic inference are constitutive of meaning, they constitute the concepts. Some theories of *a priori* justification based on meaning, unlike theories, for example, such as that of Carnap (Carnap, 1937; 1947), among others, are developed from some kind of *Conceptual Role Semantics*[5] (hereafter CRS) and emerge as alternative responses to the problems that are presented to theories based on intuition (Dogramaci, 2012). CRS theories are presented in a variety of ways, just as there are different theories of meaning.

---

[2] I will focus on this article in these Boghossian papers. His position has changed over recent years.

[3] The problem of the meaning of logical constants is closely related to that of the justification of basic logical knowledge (Gómez Torrente, 2007; MacFarlane, 2015).

[4] "In considering implicit definitions, we must bear in mind that they come in two varieties: explicit and implicit. An *explicit* implicit definition involves an explicit stipulation by a thinker that a given sentence S(f) is to be true if its ingredient term f is to mean what it does. In the implicit variety, it is somehow tacit in that person's behavior with the term f that S(f) is to be true if f is to mean what it does" (Boghossian, 2008b, chap. 10, p. 218).

[5] While conceptual role semantics (CRS) in the philosophy of language is a theory of linguistic meaning, in the philosophy of mind, it is a theory of the mental content of attitudes such as beliefs or desires (Whiting, 2009).

In their diversity, what they have in common is that they are theories of meaning-as-use whose main precursor is Wittgenstein.[6] In such theories, the meaning of an expression (or the propositional content of an attitude) is determined by the role it plays in the subject's language (or knowledge; Whiting, 2009).

One of the ways of considering CRS is the so-called inferentialist conception (Brandom, 2000), according to which, apprehending the concept or understanding a word is determined by the disposition to infer according to the schema or pattern of inference of the concept. Thus, it can be stipulated, for example, that *I* is the concept, "I" is the term and *P* the schema of inference of which *I* is a constituent part. That is, for each concept there is, according to Brandom, a scheme of inference that is constitutive of that concept; so that, he or she who has the concept of square, must be able to infer correctly: if *x* is square, then *x* is not round, where *x* is an object visible to the subject in question. Being willing to infer properly following the right pattern, requires satisfying the possession of concept *I* or term "I". On the contrary, if someone is not disposed to infer following the correct scheme, he or she does not understand the term or does not understand the concept. In some way, this approach comes to the fore in his endeavour to spell out what is involved in our command of concepts in terms of inferential abilities and our disposition to infer according to the logical rules. These aspects point to the core of the dispute between inferentialist and representationalist conceptions of language. While for inferentialists, judging has the conceptual priority since asseverative contents are the smallest units that can register in inferences, for representationalists, thinking about F (if-then propositions, for example), is basic to their theory (Fodor & Lepore, 2002). So, according to representationalists, an abstract object like a proposition has primacy over propositional thought. Inferentialist conceptions—including Boghossian's proposal, and the conceptual role semantics (CRS) under any of its forms—have received extensive criticism from Williamson (2007) and, among others, also from Fodor (one of the advocates of representationalism) and Lepore (1993).

In this article, I will first analyse Boghossian's theory of justification based on meaning. Boghossian maintains the rule-circular justifications from the conceptual role semantics position, i.e. that principles of logic, as theoretical principles, are part of the meaning of language expressions. Secondly, I will present how the justification of deductive practice can be connected with the inferentialist conception of the meaning of logical constants. Considering that Boghossian explains this concept based on the notion of blind reasoning—that is, the basic inferential competence that is held prior to any explicit belief about logical validity or the conceptual resources necessary to articulate them—we are able to engage in blind reasoning: a capability to use rules without knowing those rules. Some of these rules can establish a kind of warrant.

[6] Several philosophers uphold one or another version of this approach; some among many others are: Strawson, Sellars, Field, Harman, Block, C. McGinn, Peacocke, Brandom.

Boghossian (Boghossian, 2003b) emphasizes the distinction between inferential and constitutive construals of the relation between meaning and entitlement. Recently, Boghossian (Boghossian, 2015) stated that "rules themselves should be thought of as contents" and "contents are best thought of as abstract objects […] as numbers, properties and propositions, rather than *concreta* like tables or (token) books" (Boghossian, 2015, p. 4).

On the other hand, Williamson (2011) rejects this inferentialist explanation and offers, instead, a causal explanation. From Williamson's point of view, competence regarding public language depends on participation in linguistic practice: causal interrelations of speakers unify (in part) a conceptual or linguistic practice (Williamson, 2011, p. 504).

Second, I argue that, if basic inferences or logical rules are constitutive of meaning—that is, they constitute the logical concepts and the content of a rule is an abstract object (Boghossian, 2015)—then abstract objects—like, for example, rules—are constitutive of meaning. From this approach, some philosophers consider that these logical concepts are abstract objects and that these concepts "must have been brought into being by the creative activity of human beings" (Boghossian, 2015, p. 4). Boghossian believes that the role of conceptual creativity is understood better "within a framework in which we talk not about *creating abstracta*, but about *selecting* them, or *discovering* them" (Boghossian, 2015, p. 11).

Finally, I expound the idea that rules, conversely to Boghossian's approach, come to exist as a result of certain kinds of mental action. Such a possibility occurs when something that we call "mental action" is carried out. The role of these actions is, roughly speaking, to function as constitutive norms, in this sense, logic is constitutive of thinking; thinking, as second nature (McDowell, 1994), takes place as part of human enrichment.

1. BOGHOSSIAN'S INFERENTIALIST PROPOSAL

1.1. Basic Inference as Constitutive of Meaning of the Logical Concepts

In this section, I am going to analyse Boghossian's meaning-based theory of justification and the question of how to connect the justification of deductive practice with the inferentialist conception of the meaning of logical constants.

Boghossian explains this theory based on the notion of b l i n d r e a s o n i n g; blind reasoning is the basic inferential competence that is held before any explicit beliefs about logical validity or the conceptual resources necessary to articulate them.

Boghossian's thesis maintains that one can justify the belief that a rule is correct from the knowledge of the meaning of the expression, that is, knowledge of the meaning of the conditional, for example, is sufficient to know that *Modus Ponendo Ponens* (MPP) is a valid inference rule. Moreover, according to inferentialism, the meaning of the expression is constituted by the rules of inference. It is said of these rules that they cannot be justified in terms of any other because

they are constitutive of the meaning of the logical constant in question; in this case, the MPP is a constitutive rule of the meaning of the conditional. Logical principles admit, then, a kind of justification that is *a priori* and is rule-circular. The basic rules of inference are justified by inferences among which are those same basic rules that they are intended to justify. The problem for rule-circular justifications and inferential justification in general, in the case of logical knowledge, is that both are rejected because of their circularity.[7] For example, we consider that MPP is the only non-derivative inference rule, if we want to justify *a priori* the MPP from the knowledge of the meaning of the logical constant "if, then", in this justification the MPP is assumed in at least a step, without previously being justified (Boghossian, 2001, p. 10). That is, we acquire knowledge of the MPP using only the logical rules and no non-inferential knowledge (Wright, 2001, p. 68). In short, the problem is that in the deduction of the validity of the MPP, we use the MPP, which generates a vicious circle. Specifically, in ordinary circular justifications, one observes two mistakes; first, begging the question: it assumes what it is trying to prove. The conclusion explicitly asserts the presupposed premises stated at the beginning, which goes against the idea of what it means to prove something or argue it.

Second, bad company: if we accept the rule-circular justifications, we will be able to demonstrate the correctness of rules that are not correct. The rule-circularity allows you to prove anything, even if it is unjustifiable (Boghossian, 2001, p. 11). Since an argument justifies an assertion only if it comes from premises that are justified, a related question is whether the knowledge of the premises depends on prior knowledge of the conclusion. This kind of circularity is usually called epistemic circularity. However, Boghossian considers that the MPP can be justified inferentially by rule-circular justification, which requires explaining that not every rule-circular justification falls into these two errors.

Although circular argumentation is not inherently fallacious, it can be if the argumentation is used to conceal, in a certain way, that one fails to prove something completely. Firstly, begging the question (Hansen, 2015) is defined as an argument in which what one wants to demonstrate is presupposed. Suppose that someone is asked to prove that this book (whose authorship is the subject being discussed) was written by Gala, to which that person responds by saying: "All the books here were written by Gala". Without independent evidence, the premise that "all the books were written by Gala" considers warranted the claim that "this book was written by Gala", instead of demonstrating this by satisfying the requirements of proof. In this case, the question is whether an argument in which the MPP rule is used justifies the validity of that rule since we are using the same rule whose validity we try to prove (Boghossian, 2000, p. 248).

[7] Along with Devitt (2005), which is based on Braithwaite, Dummett does not see any problem here. This is what he calls a "pragmatic circularity" (Boghossian, 2008b, p. 199).

Boghossian argues that the argument in which the MPP rule is used to justify such a rule (and so, the conclusion is already presupposed before having proven it), would not amount to begging the question if the knowledge of the premises is obtained properly regardless of the conclusion. To solve the problem of begging the question, then, we must resolve the question of whether the justified rule-circular belief that a rule, such as MPP, preserves the truth enables us to use that rule of inference, and how a circular-rule argument can warrant its conclusion (Boghossian, 2001, p. 12).

On the other hand, we have the bad company problem. Boghossian points out that there is a big difference between an ordinary (or grossly) circular argument, where the conclusion is one of the premises, and a rule-circular argument, in which the rules whose validity is to be demonstrated are used during the proof (Boghossian, 2000, p. 245). In any case, the problematic situation that arises is that a rule-circular justification seems to be available for any rule (Wright, 2001, p. 49), that is the objection of bad company, as the connective ("tonk") of the example of Prior (1960).

In his brief article *The Runabout Inference Ticket*, Prior criticizes a certain idea of the definition of the logical constants. Prior takes the case of the conjunction to present the theory of meaning that he intended to discuss. According to that theory, the rules of introduction and elimination of the conjunction establish the meaning, and because someone knows the meaning of that constant, he or she knows how to infer using the conjunction. Similarly, the connective "tonk" could be established:

> Its meaning is completely given by the rules that (i) from any statement P we can infer any statement formed by joining P to any statement Q by "tonk" (which compound statement we hereafter describe as "the statement P-tonk-Q"), and that (ii) from any "contonk-tive" statement P-tonk-Q we can infer the contained statement Q. (Prior, 1960, p. 39)

Prior argues that the definition of logical constants is not clearly determined and that any statement can be inferred from another in an analytically valid way; for example, using the new constant "tonk" we can infer in two steps "2 plus 2 equals 5" from "2 plus 2 equals 4" (Prior, 1960, pp. 38–39). This approach of analytical validity creates a plethora of other possibilities for other rules with connectives such as "tonk", that is, it allows one to produce rules of inference arbitrarily. This is one of the objections that any inferentialist theory must solve.

Boghossian (Boghossian, 2000, p. 251, note 19), following Dummett (Dummett, 1991, p. 202), considers that not all rule-circular arguments involve the two errors indicated. He proposes a theory based on the conceptual role to deal with the problems that circularity presents. According to this theory, a genuine definition of the logical constant in question is required in such a way that the arguments based on the basic inference rules linked to that constant are valid. This definition will prevent a rule-circular justification from being available for any

rule. Defended in this way, the rule-circular arguments avoid the charge of bad company. According to Boghossian (Boghossian, 2001, p. 28), the dilemma is to explain how a person can reason in a certain way, without that person's knowing (knowing the validity of) the rule contained in his or her reasoning. Boghossian illustrates what our constants mean based on their conceptual role:

> (…) that our logical words (in the language of thought) mean what they do by virtue of their inferential role, that "if, then", for example (or more precisely, its mentalese equivalent) means what it does by virtue of participating in some inferences and not in others. If this is correct, and if, as is overwhelmingly plausible, it is by virtue of its role in fundamental (i.e., underived) inference that the conditional means what it does, then we have an immediately compelling answer to the question: how could someone be entitled to reason according to MPP without having a positive belief that entitles him to it. *If fundamental inferential dispositions fix what we mean by our words, then, as I shall now try to show, we are entitled to act on those dispositions prior to and independently of having supplied an explicit justification for them*. (Boghossian, 2001, pp. 28–29, italics mine).

Then, Boghossian proposes that for the case of basic inferences, the rule involved must be m e a n i n g - c o n s t i t u t i n g; this would explain why the person is entitled to use the rule without a demand that he or she knows that the rule preserves the truth (Boghossian, 2001, p. 29). So, he formulates the following principle:

> (L) If M is a genuinely meaning-constituting rule for S, then S is entitled to infer according to M, independently of having supplied an explicit justification for M. Notice that (L) does not require that S know that M is meaning-constituting for S if S to be entitled to infer according to M but only that M be meaning-constituting for S. (Boghossian, 2000, p. 250)

Boghossian must argue that rule-circular justifications are genuine justifications as long as we distinguish between a rule-circular and an ordinarily (grossly) circular argument, since the latter does not guarantee trivial success. If you rely on a small number of uses of a particular rule, an adequate rule-circular argument allows you to support that rule preserving the truth in any possible use necessarily only in the instance that the rule in question is constitutive of the meaning (Boghossian, 2000, p. 254). For this, it must be established that the meaning of logical constants is determined by their conceptual role. If an inferential disposition constitutes the meaning, then it is *a fortiori* reasonable, so it can be used justifiably without a supporting argument (Boghossian, 2000, p. 250).

On the other hand, one of the problems of the CRS is that it is based on a holistic approach to meaning (or content) (Whiting, 2009), this implies that an expression cannot have meaning by itself, since, as we have indicated above, an expression is significant by virtue of its inferential relations. This represents a difficulty from the point of view of communication since, according to this semantic approach, the inferential meaning of an expression will depend on the

beliefs that a subject has. If two subjects have different beliefs about an expression, each one will be willing to infer according to inferential transitions that understand that expression, and if the beliefs of the subject change, the inferential transitions, therefore, will be diverse. As Martínez-Vidal states:

> (…) intuitionists and classical logicians have tried to solve the dispute putting forward the import that maintaining one position or the other has for mathematical practice. But of course, determining whether a given mathematical practice is right depends, to a certain extent, on the philosophical thesis we assume. This is so because our intuitions or judgements about the correctness of a given argument will differ depending on (…) our philosophical views. (Martínez-Vidal, 2004, p. 204)

Hence, if we consider that the same word, according to the CRS, has different meanings and is to be understood in different ways, it will be practically impossible that the meaning remains constant (Whiting, 2009). To avoid these problems, several authors, including Boghossian, argue that analytical/synthetic distinction is needed to differentiate the inferential transitions that determine the meaning (or the content) and those that do not:

> This would provide something constant—an invariant significance—that could be grasped despite differences in belief. And, moreover, it respects compositionality, since the meaning of a complex expression is fixed only by its role in analytic inferences, and that is determined by the meaning of its parts. (Whiting, 2009)

In *Epistemic Analyticity: A Defense* (2003b), Boghossian presents his defence of epistemic analyticity to explain our knowledge of the validity of basic principles of inference. Boghossian takes the notion of analyticity as a property of linguistic items from which grasping the meaning of a sentence is enough to justify belief in the proposition it expresses. Or similarly, grasping a proposition *p* is sufficient to justify his belief in *p*:

> I will talk of grasp of the meaning of a sentence as sufficing for justified belief in the proposition it expresses; but I could equally well have talked simply about grasp of a proposition p as sufficing for justified belief in p. Thus, too, I will talk about words being synonymous with each other; but I could equally well have talked about concepts being identical to one another. Finally, I will talk of holding some sentences true, as a condition of meaning some specific proposition by them; but I could equally well have talked of believing some propositions as a condition of having some of their ingredient concepts. (Boghossian, 2008b, p. 212)

Epistemic analyticity is a way of explaining how factual propositions can be known *a priori*. The models for the construction of the epistemic analyticity that Boghossian presents are three: the Frege-analyticity, the Carnap-analyticity and the constitutive model (Boghossian, 2008b). According to the constitutive model,

the semantic facts themselves provide the necessary justification, as opposed to the knowledge that a subject can have of them (Boghossian, 2008b, pp. 4–5). Boghossian states that, by the mere fact that the subject grasps the meaning of the rule of inference from R, it implies that this subject is enabled to infer according to R. Thus, any inferential transition from the conditions of possession of a concept is by itself, *prima facie*, entitling, that is, it gives a warrant.

The problem is how to avoid the charge of bad company faced by this model that seems to have counterexamples of the type of "tonk", which we have already presented; that is, the rules for "tonk" are not correct, they allow one to move from true premises to any conclusion, including false conclusions (Boghossian, 2001, p. 13); but what is wrong in a rule-circular justification of one of the "tonk" rules is not circularity (Boghossian, 2001, pp. 33–34), but the fact that practice according to that rule fails to constitute any meaning:

> (…) a practice which allowed that "A tonk B" may be inferred from either A or B individually, and that both A and B individually may be inferred from it, would establish no meaning for "tonk". By contrast, the practice of inference in accordance with *modus ponens* is part of a meaning constituting practice: a practice which constitutes the meaning of "if…, then…" And that, ultimately, is why we may in principle justify the belief that *modus ponens* is sound by a derivation which uses *modus ponens* in its course. (Wright, 2001, p. 52)

To safeguard this model, Boghossian suggests the following: restrict the model to certain concepts in which an entitlement is given or restrict what we are going to consider as a genuine concept. This is the correct way to understand the conceptual semantic role (Boghossian, 2001) and avoid the problem of bad company.

> A conceptual role semantics, by virtue of its ties to the notion of justification, transforms this constraint on meaning into a constraint on justification that simultaneously vindicates the possibility of rule-circular justifications while staving off the threat of an unpalatable relativism. (Boghossian, 2001, pp. 33–34)

Thus, Boghossian proposes an improved version of the constitutive model: any rules written into the possession conditions for a non-defective concept are *a fortiori* entitling. So, how could a thinker be entitled to reason according to MPP just by virtue of grasping the meaning of that rule? The answer is that he or she can be so entitled because MPP is a possession condition for the conditional, and the conditional is a non-defective concept.

2. CONCEPTUAL ROLE DETERMINES A MEANING?

According to Boghossian, following certain rules of inference is constitutive of our understanding of primitive logical constants. Secondly, if certain rules of inference are constitutive of our understanding of certain concepts, then we are

entitled to them, even in the absence of any reflexively appreciable support. To deal with the problem of bad company, Boghossian presents a solution based on the CRS approach; but, from a purely theoretical foundation of meaning, to say that any possible conceptual role determines one meaning or another, as was mentioned, is not enough.

> We should insist that a conceptual role determines a meaning for an expression only if it manages to contribute in some determinate way to determining how the world would have to be if sentences involving the expression are to be true. (Boghossian, 2001, p. 33)

To conclude, the arguments Boghossian has offered demonstrate that the rule-circular argument for MPP allows the use of MPP (use to which we are enabled if the MPP is a rule that constitutes the meaning) to be able to determine that MPP necessarily preserves the truth in any possible use (Boghossian, 2000, p. 252). If the rules involved in the rule-circular justification are rules constitutive of meaning, we are entitled to use those rules of inference, independently of whether we are justified in believing that any such rule is valid. In this way, Boghossian argues that basic logical knowledge is justified inferentially and that circular-rule justifications are genuine justifications. To explain how a person is enabled to infer, without implying that the person knows something about the rule used in his inference, he responds that this entitlement naturally flows from considering the rules as meaning-constituting (Boghossian, 2000, p. 249), and proposes his approach concerning warrant transfer (Boghossian, 2001, p. 29).

On the other hand, Williamson (2007) argues that all knowledge is propositional. Thus, while Boghossian starts from the fact that linguistic or conceptual practice is a precondition for understanding certain links, such as understanding-assent, Williamson objects to Boghossian that no justification or knowledge can be derived from linguistic or conceptual competence. Among other reasons, while knowing implies assenting or accepting, from the assertion or acceptance knowledge does not necessarily follow (Williamson, 2007, p. 76).

Williamson, in addition, underlines an aspect that is important. For Boghossian, logical-term competence involves assenting to the kind of understanding-assent link required (Williamson, 2011, p. 503). Williamson argues that such links are not necessary. When a word or term belongs to a public language, being competent about that term involves causal relationships with other speakers, other subjects. The question is what explanation of the linguistic competence is pursued.

According to Boghossian conceptual practice is a precondition for understanding links such as understanding-assent, but for Williamson (2007), the idea that understanding-assent links are the case belongs to an inaccurate theory of meaning, according to which, if these links do not occur, the distinction between understanding and not understanding is dissolved. Speakers who understand the same term may have nothing in common that constitutes a shared meaning.

On the other hand, Williamson offers an alternative theory of meaning according to which such links do not exist and there is a shared language. He rejects the thesis that the shared understanding of a word requires a stock of shared obviousness. This is because to defend this presupposes acceptance that the uses by different agents can be united to form a common practice of using that word with a given meaning only due to an invariant core of beliefs. Therefore, according to Williamson, the idea that understanding is epistemologically sufficient to assent is based on a false conception of what it is to understand. The social determination of meaning is not based on the idea that meaning cannot be determined individually; but, when an individual uses a shared language as such, the individual meaning is parasitic on social meaning. Much of the practical value of language is its ability to facilitate communication between agents in asymmetric epistemic positions when the speaker knows something that the listener does not know. The practical constraint for communication is that there must be a background of broad agreement in the use of the terms. This practical constraint is holistic (the agreement at a certain point can be exchanged by agreement with others). However, the existence of a broad agreement is not a necessary condition. For example, being competent in the use of English does not require the acceptance of the principle of non-contradiction.

In this way, Williamson rejects the inferentialist explanation and offers, instead, an approach based on the causal interrelations between speakers; it is these relationships that constitute the competence and we agree with this explanation: the competence regarding public language depends on participation in linguistic practice. Causal interrelations of speakers unify (in part) a conceptual or linguistic practice (Williamson, 2011, p. 504). From my point of view, and as McDowell (1994) pointed out, a shared language is a primary medium in which understanding is generated.

On the other hand, Wright (2004a) argues that being constitutive of meaning based on b l i n d r e a s o n i n g is not sufficiently explanatory since it does not clarify how "blind" inferences confer knowledge. Thus, the problem of Boghossian's approach is not with respect to the acquisition of knowledge, as Wright contends, but with respect to the justification for it. Blind reasoning as an explanation of the second is not enough.

According to Wright, we can consider a kind of justification in a non-inferential weak sense, a rational warrant that does not require evidence of truth and can avoid both circularity and the infinite regress of justifications. This warrant consists of a mode of acceptance of a proposition and may be rational but not equivalent to belief (in the primary meaning of belief). It is the entitlement to assume the initial presuppositions, as long as there is no evidence against them, even if there is no evidence in their favour (Wright, 2004a, p. 161). This rational warrant is also a viable solution for the particular case of logic. Wright argues that basic rules of inference such as the MP are one of the types of initial presuppositions. We can trust in the validity of basic inferences, and this, in the end, enables us to state knowledge of the reasoning products obtained through the

application of that basic logic (Wright, 2004b, p. 208). In the case of logical knowledge, a logical proposition is presented with a statement as following necessarily from the premises within a given logical system.

If, in addition, one considers that basic inferences are basic mental actions that function as constitutive norms (Wright, 2014), then one accepts that the meaning of logical expressions is given by the basic rules. For Wright, justification does not derive from conceptual understanding but rather, the rules should be assumed as a starting point.

However, this approach presents several problems (Garcia-Arnaldos, 2017). On the one hand, it is not clear how to explain the status of the fundamental rules of inferences. On the other, basic inferences can assume patterns that are not very solid, sometimes we make mistakes, for example, confusing the directionality of if-then propositions (*fallacious modus ponens*). To solve these problems, we must first answer the question, what is a rule?

3. Abstract Object as Content for a Rule

Regarding how can we define a rule, Boghossian claims that "rules themselves should be thought of as contents" and "contents are best thought of as abstract objects (…) as numbers, properties and propositions, rather than with *concreta* like tables or (token) books" (2015, p. 4).

The question whether there are abstract objects such as numbers, universals, and propositions, is analysed in many contemporary philosophical debates. Several philosophers argue that abstract objects exist, but they are conceived as mind-independent objects (without causal contact). Some philosophers consider that logical concepts are abstract objects and that these concepts "must have been brought into being by the creative activity of human beings" (Boghossian, 2015, p. 4); for instance, Thomasson's conception of some abstract objects as artifacts (Thomasson, 2014), i.e. tools designed to improve our ability to represent reality. Others are not willing to accept abstract objects in their ontology.

Boghossian believes that the role of conceptual creativity is understood better "within a framework in which we talk not about *creating abstracta*, but about *selecting* them, or *discovering* them" (2015, p. 11):

> When one is born into a society that has accepted certain norms and lives by them, and if one continues to live with and benefit from that society, then, other things being equal, one is obligated to live by the norms that are accepted in that society (…).
>
> The main point right now is that it would be a mistake to look for a source of normativity either in the rule itself or in the mere fact that a rule has been accepted. If there is an obligation to obey a rule it cannot come from any source other than from the requirements of morality, which, as I previously emphasized, provide a norm on behaviour independently of whether they have been accepted. (Boghossian 2015, p. 11)

In my belief, rules come to exist as a result of certain kinds of mental action, when a certain mental action is carried out. The role of these actions is, roughly speaking, to function as constitutive norms, as Wright pointed out. The word "norm" or "normative" is usually defined as involving a rule or correctness. Normative inferentialism maintains that the meaning of a word is constituted by rules or norms governing inferences, so the meaning of expressions is constituted by the rules of inference. For Broome (2013; 2014b), it means involving a reason or ought.

My premise is that logical rules are not abstract objects, but ideal, similar to what Wittgenstein (1953, §38) called "ideal".[8] What we construe as norms or rules of content, inferential rules such as MPP, may involve idealization, but this is because we already share a language within a given culture. According to Railton:

> Logic does involve idealization, the creation of "crystalline" models. But the function of these models is not to give us an ideal for all thought and language, an image of how the content of our thinking would be structured if all were right with us (…).
>
> The *norma* and *regula* were said to function regulatively for us as builders *a priori*—standards we require our cuts to meet, and correct them to fit (…). Wittgenstein says of logic: "the crystalline purity of logic was, of course, not a result of investigation: it was a requirement" (PI 46). But what sort of requirement? One sense might be *a logico-metaphysical requirement*. (Railton 2000a, p. 189)

Boghossian assumes the source of normativity as requirements of morality (in a society). The source is not in the rule itself and it is not in the mere fact that a rule has been accepted. The question is whether—in the requirements of morality—the conceptual priority is in the judgment or in the object. Boghossian doesn't clarify this issue. I believe that the conceptual priority is in both, in the judgment and the object (in this case, the logical rules) and the requirements are logico-metaphysical ones. That is, logical rules are neither prescriptions of thinking nor psychological laws. "Their 'validity' or necessity is *sui generis*; if anything, it is what we might today call metaphysical" (Glüer & Wikforss, 2018).

These points are to be stressed; as Wright (2014) argues, one is able to rationally rely on the validity of basic inferences because they are basic mental actions that function as constitutive norms, (in the same way that one relies on deductive reasoning since one of the primary functions of reason—as Burge [1993] states—is to present truth). Then, a basic inference will be a norm of rational action, in such a way that only by the activity and in accordance with it, does a subject achieve intentionality. Furthermore, all action has a directiv-

[8] "(T)he most that can be said is that we construct ideal languages. But here the word "ideal" is liable to mislead, for it sounds as if these languages were better, more perfect, than our everyday language: and as if it took the logician to shew people at last what a proper sentence looked like" (Wittgenstein, 1953, §38).

ity, in that it is directed to something beyond the action itself. If one does not want to fall back into the regression to infinity produced by the recourse to intentionality, this can be understood as a guiding disposition.

For Broome (2014a), the mental action in which inference consists is a guiding disposition. Broome proposes that when one reasons or infers, one is guided by the rule. The guidance given in the rule is intentional guidance.[9] This intention or habit is a disposition that guides the reason: "An intention is a sort of disposition to behave in a particular way." (Broome, 2014b, p. 629). So, when one intends to comply with the rule, it is most likely that this intention is a habit. These intentions or habits are dispositions to the mental action that constitutes the rule.

On the other hand, one can make mistakes in inferring. That is, one makes an inference because it seems correct, but it might not be. Broome has maintained that, even without defining what that correction consists of, what is important is the difference itself, the fact that one can distinguish between seeming right and being right. One makes an inference because it seems correct, but it does not mean that one is not going to make a mistake. According to Broome, inferring wrongly is inferring anyway. Making an inference could be a problem if one wants to maintain certain logical normativity.

To solve the mistake problem, Broome presents two senses of normative. One can think of normativity as belonging to a correction standard (Broome, 2014a, pp. 24–25). If one conceives normativity in that way, a rule would be automatically normative, ("right" as a synonym of "correct"), but then, when one utilized correct rules, there would be no error, no mistakes. Hence, Broome holds another sense of "weaker" normativity: there are rules that do not entail that one has reasons to follow these rules; and also, that a rule can be followed without apparent reason. This other sense of normative would involve a reason or an ought (Broome, 2013; 2014b). To make it clear, Broome distinguishes between *normative guidance* and *intentional guidance*: when one thinks one should do something, one is guided in some way to act, it is a normative guide. In this case, a belief-reason is insufficient, a belief-ought is needed. Applied to the rules of reasoning, one of the problems of normative guidance, according to Broome, is the acquisition of beliefs that a reasoning rule must be followed. If one has reasons to follow the MPP (for example), the reasons must be rational. But rationality does not require following the MPP. MPP is a correct rule of reasoning, according to Broome, because rationality allows it, not because rationality requires it. That is to say, although a rule is normative in a "weak" sense, it is not necessary that it be so in a "strong" sense; one does not need normative thinking, nor a reason to believe the conclusion when one arrives at it following a rule. But,

[9] "Intentional guidance cannot be reduced to normative guidance, then. It does not need to be, because intentions provide perfectly good guidance by themselves; they require no help from normativity. On the other hand, normative guidance does need help from intentional guidance. A normative belief does not guide you directly" (Broome, 2014b, p. 630).

from Boghossian's view, one has yet to solve this question: "In general, nothing normative follows from the mere fact that one is following a rule. It all depends on the content of the rule. So, where does the normativity of obeying the law come from?" (Boghossian, 2015, p. 10).

To answer, I would say there is a kind of "normative freedom", as Railton (2000b) states; but these may not necessarily be two opposite perspectives. It could be said that the normative combines these two elements: force and freedom. To illustrate this point by which, in a broad sense one makes inferences on a daily basis (that is, without involving a specific rule): if one tells someone that one spent the afternoon reading and did so productively; would one take the calculator and do an operation to calculate exactly how many pages one read, or would one simply say one had read about x pages? The "mental calculations" and the more or less simple reasonings that one usually applies in one's daily life do not always follow precise rules. And, on the other hand, there are rules, like the basic logical rules that are correct and objective, but it does not mean that one always has reasons to follow them. The objectivity of logic is not affected, since one has started from the distinction between seeming right and right. The basic logical rules do not cease being objective and correct just because a rule does not seem good to someone. The meaning I hold of "normative freedom" has to do with the potential capacity to give reasons to the supposed norms of reason. I put forth that, similarly, Boghossian argues that "rules and rule following facts are not normative in themselves":

> Rules are themselves abstract objects: either normative propositions or instructions. Their status as norms on behavior can be explained in some cases without anything—as in the case of true moral propositions—or, in other cases, via their acceptance, either directly or indirectly. Following a rule is not in general a problem. What is a problem is explaining rule-following in cases where there is no explicit intention to conform one's behavior to a rule. Finally, rules and rule following facts are not normative in themselves. They derive what normativity they may on occasion have from the holding of some underlying moral truth. (Boghossian, 2015, p. 11)

To complete the picture of the relations between content and normativity, I would refer to the idea of a second nature. The role of mental actions is, roughly speaking, to function as constitutive norms, as has been discussed. In this sense, we can maintain that logic is constitutive of thinking, and thinking, language, culture, as a second nature (McDowell, 1994) takes place as part of human development. One of the central issues of *Mind and World* (McDowell, 1994) is how it is possible to insert freedom and normativity into the scientific image of the world. McDowell solves this issue by overthrowing the dichotomy between the logical space of reasons and the realm of laws and introducing the notion of second nature. The need for this concept is debatable, as I argue elsewhere (Garcia-Arnaldos, 2018), but McDowell points out this important aspect: it is not our mental states that determine the meaning of our

words, but it is the relationship with the world and with other subjects that does so. We understand because we have the ability to relate to one another. I perceive the rational connections, and, in this way, these links become rational to me. But reason alone cannot be understood without the ability to go beyond itself. The starting point is knowing, but our knowledge has limits. The limits of the thinkable are determined by the characteristics of our faculty of understanding. One learns to make inferences only when one has a language and the language is always shared. Even if one invented a new language or expressions, one could not conceive them except with the elements, the linguistic "rules of the game" learned socially. To learn something new, a good inference must use a process that preserves truth; but, how does one learn to infer properly and use a process that preserves the truth? Surely, one cannot always avoid error, but as one studies ideal objects within the framework of a publicly controlled dialogic practice, one must be willing to rethink the legitimacy of the allegedly rational connections that constitute the space of reasons. Culture and science are collective activities and, as our second nature (McDowell, 1994), are part of human enrichment.

4. CONCLUSION

Upon analysing if inferential articulation is sufficient to account for conceptual content and in which sense concepts are norms determining the correctness in reasoning, inferentialists place judging over other kinds of mental acts. The question is whether the conceptual priority is in the judgment or the object and what theory of content is pursued. Grasping content as a matter of knowing how a word or concept behaves in inferences is not completely explicative. The main conclusions from the above investigations are as follows:

I. Conceptual Role Semantic is not destined to play the principal role in a justification of logic or rule-circular justification. I believe that such positions are problematic. On the one hand, Boghossian's answer to how to connect the justification of deductive practice with the inferentialist conception of the meaning of logical constants based on the notion of blind reasoning—a capability to use rules without knowing those rules—is not a definitive solution.
II. I consider with Wright (2004a; 2014) that justification does not derive from conceptual understanding, but rather that the rules should be assumed as a starting point. On the other hand, Williamson (2011, p. 504) rejects the inferentialist explanation and offers, instead, an approach based on the causal interrelations between speakers; it is these relationships that constitute the competence.
III. Boghossian (2015) also faces the problem of what is a rule. Contrary to common opinion, he sustains that rules themselves can be thought of as contents and contents are abstract objects (2015, p. 4). I discuss how abstract objects—like, for example, rules—would be constitutive of meaning. Basic inferences are basic mental actions that function as constitutive norms, but log-

ical rules are neither prescriptions of thinking nor psychological laws. I contend that the conceptual priority is in both in the judgment and the object (the logical rules) and the requirements are logico-metaphysical ones. In this sense, one can maintain that logic is constitutive of thinking. But reason alone cannot be understood without the ability to go beyond itself. One learns to make inferences only when one has a language and the language is always shared. Even if one invented a new language or expressions, I cannot conceive them except with the elements, the linguistic "rules of the game" learned socially. In this sense, competence regarding language depends on participation in linguistic practice.

ACKNOWLEDGEMENTS

Thanks to Episteme Research Group, specially Concha Martínez and Xavier de Donato. If not for the generosity of their constant encouragement and the benefit of their advices, I would never have succeeded in saying much of epistemology of logic.

REFERENCES

Boghossian, P. (1996). Analyticity Reconsidered. *Noûs*, *30*(3), 360–391.

Boghossian, P. (1997). Analyticity. In: B. Hale, C. Wright (Eds.), *A Companion to the Philosophy of Language* (pp. 331–68). Oxford: Blackwell.

Boghossian, P. (2000). Knowledge of Logic. In: P. Boghossian, C. Peacocke (Eds.), *New Essays on the A Priori* (pp. 229–254). Oxford: Oxford University Press.

Boghossian, P. (2001). How Are Objective Epistemic Reasons Possible? *Philosophical Studies*, *106*(1/2), 1–40.

Boghossian, P. (2003a). Blind Reasoning. *Proceedings of the Aristotelian Society*, *Supplementary Volume*, *77*(1), 225–248.

Boghossian, P. (2003b). Epistemic Analyticity: A Defence. *Grazer Philosophische Studien*, *66*(1), 15–35.

Boghossian, P. (2004). Is Meaning Normative? In: C. Nimtz, A. Beckermann (Eds.), *Philosophy-Science-Scientific Philosophy* (pp. 205–218), Paderborn: Mentis: PP.

Boghossian, P. (2008a). Epistemic rules. *Journal of Philosophy*, *105*(9), 472–500.

Boghossian, P. (2008b). *Content and Justification: Philosophical Papers*. Oxford: Oxford University Press.

Boghossian, P. (2015). Rules, Norms and Principles: A Conceptual Framework. In: M. Araszkiewicz, P. Banaś, T. Gizbert-Studnicki, K. Płeszka (Eds.), *Problems of Normativity, Rules and Rule-Following* (pp. 3–13). London: Springer.

Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge: Harvard University Press.

Broome, J. (2013). *Rationality Through Reasoning*. Oxford: Blackwell.
Broome, J. (2014a). Comments on Boghossian. *Philosophical Studies*, *169*(1), 19–25.
Broome, J. (2014b). Normativity in Reasoning. *Pacific Philosophical Quarterly*, *95*(4), 622–633.
Burge, T. (1993). Content Preservation. *The Philosophical Review*, *102*(4), 457–488.
Carnap, R. (1937). *The Logical Syntax of Language*. London: Routledge.
Carnap, R. (1947/1956). *Meaning and Necessity: A Study in Semantics and Modal Logic*, Chicago: Chicago University Press.
Carroll, L. (1895). What the tortoise said to Achilles. *Mind*, *4*(14), 278–280.
Devitt, M. (2005). There is no *A Priori*. In: E. Sosa, M. Steup (Eds.), *Contemporary Debates in Epistemology* (pp. 105–115). Cambridge: Blackwell Publishers.
Dogramaci, S. (2012). Apriority. In: D. G. Fara., G. Russell (Eds.), *The Routledge Companion to Philosophy of Language* (pp. 768–782). New York: Routledge.
Dummett, M. (1991). *The Logical Basis of Metaphysics*. Cambridge, MA: Harvard University Press.
Field, H. (1977). Logic, Meaning, and Conceptual Role. *Journal of Philosophy*, *74*(7), 379–409.
Field, H. (2000). Apriority as an Evaluative Notion. In P. Boghossian, C. Peacocke (Eds.), *New Essays on the A Priori* (pp. 117–149). Oxford: Oxford University Press.
Fodor, J. A., LePore, E. (2002). *The Compositionality Papers*. Oxford: Oxford University Press.
Fodor, J. A., LePore, E. (1993). Why Meaning (Probably) Isn't Conceptual Role, *Philosophical Issues*, *3* Science and Knowledge, 15–35.
Frapolli, M. J. (Ed.) (2007). *Filosofía de la lógica*. Madrid: Tecnos.
Frege, G. (1918). The Thought. In: P. Geach, R. H. Stoothoff (Trans.), *Gottlob Frege* (pp. 351–372). Oxford: Basil Blackwell.
García-Arnaldos, M. D. (2017). Elizabeth Anscombe: razones y acciones. In: M. Ríos Guardiola, et al. (Ed.), *Mujeres con luz* (pp. 89-108). Murcia: EDITUM.
García-Arnaldos, M. D. (2018). *La relación mente-mundo en McDowell. Conocimiento y experiencia*. Madrid: Ápeiron Ediciones. Colección Faber & Sapiens.
Glüer, K., Wikforss, Å. (2018). The Normativity of Meaning and Content. *The Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/spr2018/entries/meaning-normativity/
Gómez Torrente, M. (2007). Constantes lógicas. In: M. J. Frapolli (Ed.) *Filosofía de la lógica* (pp. 179–206). Madrid: Tecnos.
Hansen, H. (2015). Fallacies. *The Stanford Encyclopedia of Philosophy*. Retrieved from: http://plato.stanford.edu/archives/sum2015/entries/fallacies/
Harman, G. (1973). *Thought*. Princeton: Princeton University Press.

Harman, G. (1986). *Change in View*. Cambridge: MIT Press.
Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.
MacFarlane, J. (2015). Logical Constants, *The Stanford Encyclopedia of Philosophy*. Retrieved from: http://plato.stanford.edu/archives/fall2015/entries/logical-constants/
Margolis, E., Laurence, S. (2001). Boghossian on Analyticity. *Analysis*, *61*(4), 293–302.
Martínez Vidal, C. (2004). Normativity and Its Vindication: The Case of Logic. *Theoria*, *50*(19/2), 191–206.
McDowell, J. (1994). *Mind and World*. Cambridge: Harvard University Press.
McGee, V. (1985). A Counterexample to *Modus Ponens*. *The Journal of Philosophy*, *82*(9), 462–471.
Peacocke, C. (1999). *Being Known*. Oxford: Oxford University Press.
Prior, A. (1969). The Runabout Inference-Ticket. *Analysis*, *21*(2), 129–31.
Quine, W. V. O. (1951). Two Dogmas of Empiricism. *Philosophical Review*, *60*(1), 20–41.
Quine, W. V. O. (1953). *From a Logical Point of View*. Cambridge, Mass.: Harvard University Press.
Quine, W. V. O. (1963). Carnap and Logical Truth. In: W. V. O. Quine (Ed.), *The Ways of Paradox and Other Essays*. New York: Random House.
Railton, P. (2000a). *A Priori* Rules: Wittgenstein on the Normativity of Logic. In: P. Boghossian, C. Peacocke (Eds.), *New Essays on the A Priori* (pp. 170–96). Oxford: Oxford University Press.
Railton, P. (2000b). Normative Force and Normative Freedom. In: J. Dancy (Ed.), *Normativity. Ratio Special Issues* (pp. 1–33). Hoboken: Wiley-Blackwell.
Thomasson, A. L. (2014). *Ontology Made Easy*. Oxford: Oxford University Press.
Whiting, D. (2009). Conceptual Role Semantics. *The Internet Encyclopedia of Philosophy*. Retrieved from: http://www.iep.utm.edu/conc-rol
Williamson, T. (2003). Understanding and Inference, *Proceedings of the Aristotelian Society, Supplementary Volume*, *77*(1), 249–293.
Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.
Williamson, T. (2011). Reply to Boghossian. *Philosophy and Phenomenological Research*, *82*(2), 498–506.
Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.
Wright, C. (2001). On Basic Logical Knowledge. *Philosophical Studies*, *106*(1/2), 41–85.
Wright, C. (2004a). Intuition, Entitlement and the Epistemology of Logical Laws. *Dialectica*, *58*(1), 155–175.
Wright, C. (2004b). Warrant for Nothing (And Foundations for Free)? *The Aristotelian Society, Supplementary Volume*, *78*(1), 167–212.
Wright, C. (2014). Comments on Paul Boghossian, *What Is Inference?*. *Philosophical Studies*, *169*(1), 27–37.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 49–57
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.04
Article

JAKUB RUDNICKI *

THE LIAR, CONTEXTUALISM, AND THE STALNAKERIAN VIEW OF CONTEXT[1]

SUMMARY: My aim in this paper is to amend the Stalnakerian view of context in such a way that it can allow for an adequate treatment of a contextualist position regarding the Liar Paradox. I discuss Glanzberg's contextualism and the reason why his position cannot be encompassed by the Stalnakerian view, as it is normally construed. Finally, I introduce the phenomenon I call "semantic dissonance", followed by a mechanism accommodating the Stalnakerian view to the demands of Glanzberg's contextualism.

KEYWORDS: contextualism, Stalnaker, liar paradox, context, semantics, Glanzberg.

INTRODUCTION

There are several contextualist responses to the threat caused by the Liar Paradox (LP).[2] One of the prominent proponents of this position, Glanzberg (2001), suggests that we need to thoroughly redesign the Stalnakerian view of context to be able to give an account of the context shift that is supposed to occur in (LP).

What I want to argue for is that no such drastic modifications are needed. The plan for this paper is as follows. First, I roughly present (LP) and lay out the generic contextualist solution to it. Then, I discuss some details of Glanzberg's position, with the special stress on his motivations for rejecting the Stalnakerian

* University of Warsaw, Institute of Philosophy. E-mail: j.m.rudnicki@uw.edu.pl. ORCID: 0000-0003-0222-9370.

[1] Work on this article was funded by a National Science Center, Poland, grant under award number 2018/29/B/HS1/01868.

[2] E.g. the work of Parsons (1974), Simmons (1993).

view. Finally, I present my suggestions as to how the Stalnakerian view can be adapted to fulfill the demands of Glanzberg-like contextualism.

THE LIAR PARADOX

The standard version of (LP) rests on an observation that it is possible to formulate a sentence that says of itself that it is not true: *This sentence is not true*. It is easy to see that when assumed to be true, it comes out not true; and when assumed not true, it comes out true. This is a dreadful contradiction, a philosopher concludes. More formally:[3]

(1) (L): (L) is not true.

(2) Suppose (L) is true.

(3) Given what (L) says, if it is true (from (2)), it is not true.

(4) So, (L) is not true. ↯

(5) Suppose (L) is not true.

(6) Given what (L) says, if it is not true (from (5)), it is true.

(7) So, (L) is true. ↯

(8) So, supposing either (a) that (L) is true or (b) that (L) is not true leads to a contradiction (from (4) and (7)).

The generic approach that contextualists take on how to deal with the above contradiction involves what is sometimes called the "Chrysippus intuition".[4] This intuition captures what are supposed to be plausible truth valuations of (L), and of the judgment expressing the ascription of the truth value to (L). As far as (L), as we saw, leads to a contradiction, it should be judged as not true. On the other hand, a judgment assigning the truth value of non-true to (L) seems to be perfectly legitimate, even though it takes (L) itself to express it. Imagine the following pair of utterances:[5]

(A) [uttered at t_1 by Obama] "What Obama says at t_1 is not true."

(B) [uttered at some later t_2, after hearing Obama's (A), by Clinton] "What Obama says[6] at t_1 is not true."

[3] This analysis is a simplified version of Gauker's (2006, p. 395).
[4] The label comes from Gupta (2001).
[5] Based on Newhard's (2009, pp 345–346).
[6] Nothing hinges here on the use of the present tense in (B).

(A) is (L)'s look-alike. This means that it is paradoxical and, therefore, not true. (B), though, assessing (A)'s truth value correctly, as not true, is true. The puzzling detail about this plausible truth valuation is that both utterances seem to be saying the same thing. Therefore, the contextualist concludes, some context shift must have happened after the utterance of (A) that allowed for the difference in truth value of propositions expressed by two tokens of the same sentence. Note that this explanation is similar to the one employed for explaining how "I am the president in April 2016" is true when uttered by Obama, but not by Clinton.

GLANZBERG'S CONTEXTUALISM

Glanzberg shares the view that there must be some kind of subtle context change in cases like the one described. What follows from this is that sentences like (L) are context-sensitive (Glanzberg, 2001, p. 230). The general idea driving his investigation is that the problem of (LP) must be a problem related not to truth alone, but to truth and truth bearers, which he takes to be propositions resulting from utterances in contexts (Glanzberg, 2001, p. 226).

As far as Glanzberg involves contexts, sentences and propositions in his picture of the liar, he is willing to reformulate (L). Now what it is really saying is that it does not express a proposition that is true (Glanzberg, 2001, p. 228). Given this, we can construct an analogous paradox again. We start with the new version of (L), call it (LL), which is the sentence saying of itself that it does not express a true proposition. Then, we assume that it expresses a proposition. If that proposition is true, then, it contradicts what (LL) says. Therefore, the proposition that (LL) expresses is not true. This leads to a contradiction. On the other hand, if that proposition is not true, then it is in agreement with what (LL) says and makes the proposition expressed by (LL) true, which also leads to a contradiction (Glanzberg, 2001, p. 228). What needs to be noted at this point is that despite how it may seem, the above reasoning has not led to a global contradiction yet. The result is just the falsification of the assumption that (LL) expresses a proposition. What we need in order to retrieve our paradox, known from the Obama and Clinton example, is to push our result a bit further (note that the current stage is the counterpart of (B); in other words, at this very moment we conclude that (A) and (LL) respectively cannot be saying anything true). To do so, it is enough to realize that from (LL)'s non-expression of a proposition (i.e. the result that we already have), it follows that it does not express a true proposition, either. A cautious reader will see that this new result is just the same thing that (LL) says (Glanzberg, 2001, p. 229). More formally, the reasoning looks as follows:[7]

[7] For a slightly simpler presentation of this reasoning see the work of Simmons (2018, pp. 771–772). A more formal one is available in Sagi (2016, pp. 922–923).

(1') (LL): (LL) does not express a true proposition.

(2') Suppose (LL) expresses a proposition.

(3') Suppose (LL) expresses a true proposition.

(4') Given what (LL) says, if it expresses a true proposition (from (3')), it does not express a true proposition.

(5') So, (LL) does not express a true proposition. ↯

(6') Suppose (LL) expresses a proposition that is not true.

(7') Given what (LL) says, if it expresses a proposition that is not true (form (6'), it does not express a proposition that is not true.

(8') So, (LL) does not express a proposition that is not true. ↯

(9') So, supposing either (a) that (LL) expresses a true proposition or (b) that (LL) expresses a proposition that is not true leads to a contradiction (from (5') and (8')).

(10') So, (LL) does not express a proposition (from (2') and (9')).

(11') So, (LL) does not express a true proposition (from (10')).

(12') (11') = (LL), so (LL) expresses a true proposition.

So now, what we get by conducting a chain of good inferences based on good premises must itself be true. This means that our reasoning shows that (LL) both expresses (at the very last stage, when we realize that what we have arrived at is exactly what (LL) says; this is the counterpart of concluding that (B) is true) and does not express a true proposition (in the penultimate step, when we get that (LL) does not express a true proposition since it does not express a proposition at all; Glanzberg, 2001, p. 229). This is the very contradiction that Glanzberg's efforts aim to resolve.[8]

He does so by allowing (LL) to be context-dependent.[9] This possibly permits us to say that there is no contradiction in one sentence's expression of a true proposition and non-expression of it, since whether it does express a true proposition depends on the context in which it is uttered.

[8] I am going to follow Gauker (2006, p. 402) in thinking that this is the correct place to locate the context shift, contrary to what Glanzberg (2001, p. 233) suggests.

[9] This, of course, requires a new reformulation of (LL). Glanzberg does it by adding the world parameter as a *relatum* for the relation of expressing a proposition and a context set parameter relative to which the proposition is to be expressed. For details, see (Glanzberg, 2001, pp. 236–237).

GLANZBERG AND THE STALNAKERIAN VIEW OF CONTEXT

As plausible as it is, this idea needs to be accompanied by some explanation of how the context shift—needed by the contextualist—may occur. The context shift has to be such that the first context forbids (LL)[10] to express a true proposition, and the new one, after the shift, allows for it. The natural candidate becomes the Stalnakerian view, whose core idea is that context changes with every successful assertion made during a conversation.

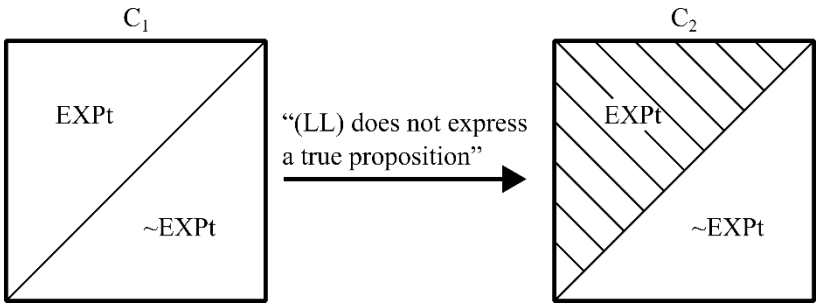
Stalnaker's position rests on the following pillars. Propositions are sets of possible worlds. Context is a set of all propositions commonly presupposed by the parties in a conversation. That makes context a set consisting of the intersection of sets of worlds. What a new assertion does, is that it divides the context set into two subsets. One of them consists of those worlds of the common ground in which the asserted content holds, and the second of those in which it does not. When the asserted proposition becomes accepted by other members of the conversation, the context set, and therefore context as such, is updated in such a way that the latter subset is eliminated (Stalnaker, 1978).

Let's see how this framework can handle the liar case. Imagine a random conversation with a random common ground represented by the context set C_1. When the reasoning leads the conversationalists to asserting that (LL) does not express a true proposition, C_1 is divided into two subsets, one with worlds in which (LL) in fact does not express a true proposition, and the second, with those in which it does. After the utterance is accepted, the context set is updated by wiping out the latter subset. In the new context C_2, it is a part of the common ground that (LL) does not express a true proposition, and the context set includes only those worlds in which that is the case. When, then, at C_2 the conversationalists want to make the Clinton-like statement, we seem to encounter a problem. (LL), when uttered at C_2, contrary to our Chrysippus intuition, cannot be expressing a true proposition, because there are no worlds in the context set in which this would be true. In other words, there are no true propositions left to be expressed by (LL) (as uttered at C_2) in our domain. A little diagram will make the situation more vivid:[11]

[10] Or its modified, contextualist, version.

[11] EXPt stands for "(LL) expresses a true proposition".

Diagram 1



As we see, when at C_2 we are to conclude, after realizing that it follows from what was uttered at C_1, that (LL) does not express a true proposition, and that it is exactly what (LL) says, it is incompatible with the set of worlds available to us for it to be the case. If (LL) now, in accordance with the Chrysippus intuition (and Glanzberg's reasoning), is to be true, it must be expressing a true proposition at C_2. As the Stalnakerian framework cannot give an account of how that situation could possibly occur, Glanzberg concludes that we ought to look for the right account somewhere else (2001, p. 239).

The main reason why Glanzberg does not see a chance of mending the Stalnakerian view is that after asserting (LL) at C_1, we simply run out of worlds in which (LL) might express a true proposition at C_2. Why the framework is malfunctioning is that it does not allow for any new source of worlds. It is monotonic in the sense, that it only allows the set of worlds to be narrowed down as we progress in a conversation (Glanzberg, 2001, p. 247).

We do not want to become entrenched in too much detail about Glanzberg's idea. Suffice it to say that, according to him, the needed source of new worlds comes from our ascent in a hierarchy of contexts. Every step higher in the hierarchy offers us more truth conditions to choose from, and therefore more expressive power for our utterances. So after (LL) is uttered at C_1, for it to be expressing a true proposition at C_2 this new context must have a bigger expressive capability.

HOW TO FIX THE STALNAKERIAN FRAMEWORK

Now, in the final section of the paper, I offer my simple way of amending the framework. I would like to focus our attention on what happens around C_2.

In my opinion, the crucial thing to be noticed is that members of the conversation "realize" that what is uttered at C_1 becomes a part of the common ground (Stalnaker, 1998, p. 99). This must result in what I call a semantic dissonance—i.e. a situation of incongruity between the set of worlds they updated to and the propositions present in the common ground. In our particular case, the semantic dissonance occurs because all the worlds that are left in C_2 are worlds

in which (LL) does not express a true proposition, but at the same time, the members of the conversation become aware that (LL) is a part of their common ground so that they must be taking its truth for granted. This last phase of rehabilitation[12] of (LL) is exactly what makes the Chrysippus intuition so plausible.[13] This, I say, suggests that there must be some mechanism that can deal with these kinds of situations when we need to make the content of propositions belonging to the common ground and the worlds available to us coherent with one another.

I consider that this mechanism can be best expressed by the following rule:

If you realize that the content of a proposition correctly added to the common ground causes an incompatibility between the contents of propositions constituting the common ground and the possible worlds available in a conversation, retract the update caused by this assertion and accommodate the context set accordingly to what the contents of the propositions in the common ground say.

The first thing worth noting is that the rule is not triggered by ordinary cases in which a certain proposition is in the common ground and the members of the conversation attempt to update it with this proposition's negation. The reason for this is that in such situations, the inconsistency between the proposition's negation and the available worlds already exists before the potential, *incorrect* update. In other words, such an incorrect update does not cause the incompatibility between the contents of the propositions in the common ground and the worlds available in the conversation, but it is this very incompatibility that makes such an update incorrect.[14]

Let's see how this rule works when put to use in the case of (LL).[15] At C_1 it is uttered that (LL) does not express a true proposition because of its semantic defectiveness. The new context set, C_2, is such that (a) it is a part of the common ground that (LL) is semantically defective and (b) it contains only those worlds in which (LL) does not express a true proposition. The participants in the conversation realize, though, that there is an incompatibility between the worlds they have left, and what the propositions in their common ground say (remember that they take elements of the common ground for granted, and therefore assume their veracity; at the same time, they do not have access to the worlds in which (LL) is true). This is what triggers our *rule*. What rational conversationalists do in such a situation is to retract the latest context update. In other words, they go to a context, in which both types of worlds, those in which (LL) is true, and those in

[12] I borrow the name from (Simmons, 2018).

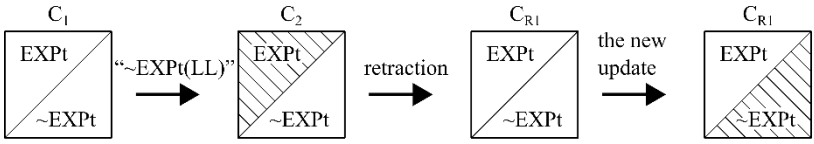
[13] This is of course not to say that the members of the conversation think of their situation exactly in these technical terms. The relevant psychological equivalent is their realization that the truth valuation of (LL) has changed. (I would like to thank an anonymous referee for pointing out the need for this clarification).

[14] The need for clarifying this issue was also signaled by an anonymous referee.

[15] I found Simmons (2018, pp. 756–760) very helpful in formulating this part properly.

which it isn't, are still accessible (let's call this resulting stage the *rehabilitation stage*). It is very important to note that this new context is not quite the old C_1. The difference between this new context (let's call it C_{R1}) and the old C_1 is that in the former, contrary to the latter, the members of our conversation already know that (LL) is semantically defective. That allows the rehabilitation to occur in C_{R1} and leads to the second update. The second update assumes (LL) to be rehabilitated (at C_{R1}) and, therefore, the new context is not C_2 but rather some C_2'. C_2' is such that it only consists of worlds in which (LL) expresses a true proposition since this is exactly what follows on from what (LL) says when reflected upon in the new circumstances. Again, I think that a diagram would make this more perspicuous to the reader:

Diagram 2



At C_2 it is already a part of the common ground that (LL) is semantically defective. The set of worlds is updated accordingly. The members of the conversation realize the inconsistency between the worlds available and what (LL) says, which leads to the semantic dissonance, triggers the rule and causes retraction.

The rehabilitation stage

The context resulting from exercising the rule and the rehabilitation of (LL).

Note again that the relevant change in context that allows the change in (LL)'s truth valuation is that in C_{R1}, contrary to C_1, it is already a part of the common ground that (LL) is semantically defective.

As may be inferred, we do not need to change the background domain of worlds to reach the true proposition expressed by (LL). This means that we do not need to employ any kind of hierarchy of contexts to make sense of the context shift that the contextualist needs to make her point. All we need is a plausible rule guiding our behavior in cases of semantic dissonance. This rule allows us to make a step backwards and retract the context update, which permits us to make sense of our Chrysippus intuition.

REFERENCES

Gauker, Ch. (2006). Against Stepping Back: A Critique of Contextualist Approaches to the Semantic Paradoxes. *Journal of Philosophical Logic*, *35*(4), 393–422.

Glanzberg, M. (2001). The Liar in Context. *Philosophical Studies*, *103*(3), 217–251.

Gupta, A. (2001). Truth. In: L. Goble (Ed.), *The Blackwell Guide to Philosophical Logic* (pp. 90–114). Oxford: Blackwell Publishers.

Newhard, J. (2009). The Chrysippus Intuition and Contextual Theories of Truth. *Philosophical Studies*, *142*(3), 345–352.

Parsons, Ch. (1974). The Liar Paradox. *Journal of Philosophical Logic*, *3*(4), 381–412.

Sagi, G. (2016). Contextualism, Relativism and the Liar. *Erkenntnis*, *82*(4), 913–928.

Simmons, K. (1993). *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge University Press.

Simmons, K. (2018). Contextual Theories of Truth and Paradox. In: M. Glanzberg (Ed.), *The Oxford Handbook of Truth* (pp. 755–786). Oxford: Oxford University Press.

Stalnaker, R. (1999). *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press.

Stalnaker, R. (1978). Assertion. In: R. Stalnaker (Ed.), *Context and Content* (pp. 78–95). Oxford: Oxford University Press.

Stalnaker, R. (1998). On the Representation of Context. In: R. Stalnaker (Ed.), *Context and Content* (pp. 96–114). Oxford: Oxford University Press.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 59–82
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.05
Article

İSKENDER TAŞDELEN *

EPISTEMO-SEMANTIC COHERENTISM: AN ATTITUDINAL VIEW OF MEANING BASED ON EPISTEMIC PRAGMATISM

SUMMARY: This article develops a conception of linguistic meaning that treats it as an attitude on the part of language users towards pairs of expressions. As with propositional attitudes, these meaning attitudes are subject to being deliberately altered over time by language users, with the aim of maximizing the efficiency of their language use. Therefore, meaning attitudes can be justified or refuted in practical terms. Our instrumentalist-coherentist approach, which allows for meanings to be advocated for alongside beliefs, provides a viable theory of justification of that kind. This view fits better with the evolutionary nature of linguistic phenomena, and resolves the problem of substitutability in opaque contexts.

KEYWORDS: coherence, justification, meaning, belief, strategic and definitory rules.

1. MEANING OF EXPRESSIONS AND THE ATTITUDINAL VIEW OF MEANING

Theories of knowledge have certainly come a long way since the time when Dewey sought to highlight the ineffective ways in which terms in this field were being used (Dewey & Bentley, 1945, pp. 225–226). From that time on, epistemologists have made progress by interpreting at least some of their preferred basic terms in the light of relations to others that they take to be more closely tied to readily accessible observations. A case in point would be Peirce's revival of the empiricist interpretation of the term "belief" as disposition to behave

* Anadolu University, Department of Philosophy. E-mail: itasdelen@anadolu.edu.tr. ORCID: 0000-0002-4248-0300.

(Misak, 2014, p. 29). As such, beliefs become testable on the basis of their actual and anticipated practical consequences, so that we can decide whether or not we should believe a proposition to be true in accordance with the pragmatist criterion of truth.

Dewey's three desiderata for *firm namings* in epistemology (Dewey & Bentley, 1945, p. 226) would therefore seem apt for generalization to cover all fields of philosophical study. Pursuing that thought a little further, we may restate them in the following way:

a) We should base our terms on publicly accessible and attainable observations.

b) We should consider the reports based on such observations as being tentative, postulational and hypothetical ones, and weigh the justificatory status of the relevant observations accordingly.

c) The terminology and the domain of observation should be chosen in a way that promotes further observations and advances.

Let us call these three Deweyan principles, respectively, *accessibility*, *hypotheticality* and *progressivity*. Adopting these as furnishing the main methodological tenets of pragmatism with regard to the philosophical clarification of theoretical terms, I wish to suggest some steps to be taken towards developing a general pragmatist view of meaning.

The first thing we should note in this regard is that Peirce's game-theoretical ideas provide another framework for developing a pragmatist view of meaning. Indeed, as is well known, the idea of extending game-theoretical semantics to natural language has evolved into a lively research program. Lewis (1969) contributed the pioneering work in this field, while more recently, Pietarinen (2014) has presented a development of Peirce's and Hintikka's game-theoretical ideas that elucidates their relevance to meaning in natural languages with reference to Grice's theory of meaning. Pietarinen (2006) gives a detailed presentation of Peirce's pragmatic view of meaning, and an interpretation in the light of both game-theoretical ideas and Hintikka's own conception of logic. In principle, the present view does not stand in opposition to work in that area: rather than being construed as providing an alternative model of the interpretation of strategic meanings to that offered by game-theoretic semantics, it can just be interpreted as providing an additional conceptualization of *strategic meanings*—one that is in line with the general ideas of Dewey. As I hint in Sections 2 and 4, epistemo-semantic coherentism can be vindicated from a game-theoretical point of view as much as from a modal or probabilistic one, in that the conditions of coherence can be revised accordingly. (This is mainly accomplished by replacing the notion of inductive support with that of the game-theoretical solution.)

Dewey's principles direct us immediately to put aside many widely supported views of meaning, such as those encountered in conceptualist and possible-

worlds-based theoretical approaches, *in so far as* concepts or possible worlds are themselves regarded as amounting to undefined metaphysical entities. For in this context, their terms will not in fact be based on accessible observations. On the other hand, attempts to define the meanings of particular expressions in terms of innocuous entities are most likely to fail, for the reason that these entities have definite characteristics that meanings of expressions cannot have. This is certainly the case for most behaviourist and physicalist explanations. Similar remarks apply, for example, to sets, types or other abstract entities, due to discrepancies between the basic properties and relations of these entities and those of meanings themselves.

Some recent views—most notably those of Horwich (1998; 2005), Grice (1957; 1975; 1978), Searle (1969; 1983) and Brandom (1994)—do not leave the linguistic realm in search of meanings, but rather appeal to regularities of use, linguistic intentions, or pragmatic inferences. These views claim that every particular semantic fact concerning the meaning of a linguistic expression can be derived from some other linguistic fact or facts. As such, they are often criticized for being circular: the objection is that meanings, if held to be explanatorily required at all, are so only because it is thought that they are needed to explain phenomena that themselves involve the use of language, such as our inferences, rather than the other way around. This sort of circularity objection against inferentialism has been raised by several authors (e.g., Fodor & Lepore, 2001). The strongest inferentialist response (Peregrin, 2009, pp. 168–171) seems to be the assertion that content is dependent on normative attitudes, but it is hard to see how normative attitudes could be explained on the basis of inferences without first giving an account of meaning.[1] Insofar as the inferentialist thesis just is the assertion that the meaning of an expression is its inferential role, (where the inferential role of an expression is determined by the inferential rules governing the use of that expression within some inferences), it is hard to see how it could overcome the circularity objection. If, on the other hand, the inferentialist thesis is that the inferential roles of expressions with respect to analytic inferences and meanings of expressions coincide, then it becomes a less interesting thesis—one that merely affirms the extensional equivalence of meaningfulness on the one hand and possession an inferential role on the other, while refraining from giving any account of meaning that could explain *why* inferential roles and meanings of expressions coincide (in the sense of two terms with the same inferential role being synonymous).[2]

[1] See the work of Sellars (1948; 1953), Grice (1957) and Brandom (1994). See also (Goble, 1967) for a further elaboration of Sellars' view and see e.g., (Macbeth, 2010) for a critical discussion of the views of Sellars and Brandom.

[2] Szubka also argues for the conclusion that Brandom's view "faces the dilemma of being an unhelpful platitudinous doctrine or theoretically fruitful but implausible conception requiring the reducibility of semantic categories to narrowly conceived pragmatic ones" (2010, p. 173).

Neither of the above remarks should be considered an attempt to furnish an overall argument against the aforementioned foundational theories of meaning. An attempt in that direction would well exceed the scope of the present paper. However, I do view these as providing an adequate motivation for refraining from trying to identify meanings of particular expressions as extralinguistic entities, or derive particular facts about meanings of particular linguistic expressions from facts that are held to be describable without invoking the notion of meaning. Instead, I propose that we consider meanings to be our irreducible collective attitudes directed primarily towards pairs of expressions: more precisely, we should think of a meaning attitude as being directed towards a pair of expressions (considered as a pair of types of physical entity consisting of actual or possible utterences of the types in question) as its object. On this account, our possession of a meaning attitude necessarily involves our entertaining an intention to revise our language so that we use the two expressions interchangeably, and our expectation that our language will function more effectively if we decide to deploy utterances of these two expressions interchangeably. The satisfaction of the latter involves the realization of such an amendment to our language, as well our future observation that our revised language is now functioning properly and would not be functioning so well if we had left it as it was, or developed it in some other way.[3]

Our meaning attitudes do not directly relate our words to the world: we only require that expressions paired by our meaning attitudes also be co-referential. The direct linking of our language with the world is achieved by means of the existence of expressions that refer directly to things (be they linguistic or non-linguistic), with reference explained in terms of the causal theory of reference. The main function of our meaning attitudes, once justified in combination with our beliefs, is to support analytic inferences within our belief systems. It follows that our proposed conception of meaning is one that aims to furnish an explanation of the semantic processing of the expressions required when making analytic inferences, where such an explanation is accomplished by an appeal to our pragmatically testable collective attitudes.

We normally express our meaning attitudes through collective assertions of the form "e_1 means e_2", where e_1 and e_2 denote two expressions of our language. These collective meaning assertions should be considered hypotheses constructed mainly on the basis of our projection of a more effective language, rather than logical or metaphysical conclusions pertaining to a prior delimited set of facts (including facts about our prior usage of the relevant terms, such as previous inferences that are in some sense valid).

Since it explicates meaning assertions as hypotheses, the present attitudinal view of meaning requires that the semantic endeavour should be carried out

[3] I have adopted this characterization of meaning attitudes from the general explanation of individual propositional attitudes developed by Vanderveken in several of the latter's writings (see, for example, Vanderveken, 2011, sect. 1).

largely within the context of the justification, rather than the discovery or explanation, of meanings. Firstly, a meaning attitude of a linguistic community should be such that either it has already been expressed by being collectively asserted, or the linguistic community is ready to accept this meaning assertion in the light of an inquiry, say, in response to a learner of the language who is competent enough to ask what some expression *e* of the language means. (Nevertheless, on the present view, the answer will be interpreted not as giving the meaning of *e*, but rather as expressing an attitude towards two expressions, namely *e*, and another expression, *e′*). Therefore, from the perspective of the present position regarding meanings, given a reasonably rich language, we need not bother about *how* to deal with questions concerning the meanings of expressions.

Secondly, the question "What are meanings?" should also not raise difficulties: if one has an understanding of propositional attitudes, such as beliefs, then a grasp of meanings should not be far away, insofar as meanings are attitudes towards pairs of expressions. Thus, the justification of meanings, rather than their discovery or explanation, becomes the central question to be addressed when seeking to explore theories of meaning of the sort under consideration here. This runs parallel with what has been observed in the field of epistemology, where neither the discovery of one's actual set of beliefs, nor the explanatory question "what are beliefs?", seem as central as issues pertaining to the formation, justification and revision of beliefs. My aim, then, is to defend the position that holds that, through utilizing the close interaction of our meanings and beliefs, it is possible to develop a viable theory of the justification of meanings as a more or less straightforward extension of the idea of epistemic coherentism.

Sellars' and Brandom's inferentialist views of meaning are also sometimes characterized as coherentist. Recall that according to their inferentialist stance, the meaning of a linguistic expression just is its inferential role. (They mainly differ on which set of inferences should be considered.) The inferential coherence of our beliefs is held to be a basic virtue of belief systems, as given an inferentially coherent belief system, it becomes possible to identify meaning as inferential role in that system. Yet this being the case, continuity with epistemic coherentism breaks down. The reason is that epistemic coherentism does not primarily aim at identifying what an agent's beliefs are, or defining the notion of belief; rather it is a methodological program aimed at identifying which belief systems should be ranked more favourably than others.

Before pursuing the task of further elaborating the position outlined above, it is worth noting that meanings as attitudes towards pairs of linguistic expressions may be viewed as providing an adequate basis for a philosophical theory, inasmuch as it permits us to formulate both major philosophical problems about meaning, and our solutions to these, within the framework of the proposed view. Among these problems, the following two are especially important, as meanings are thought to play an indispensable role in addressing them: firstly, reference

fixing for the terms we use, and secondly, making explanatory sense of the existence of informative identity statements.

As regards the former, we may note that it is by no means universally accepted that the meaning of an expression considered as an entity can help us explain how we could fix the reference of that expression. On the contrary, there is actually a strong element of doubt as to whether meanings could ever do that. Moreover, in the context of some version of the Kripkean causal theory of reference, our terms can refer without the mediation of meanings. Assuming such a theory, co-referentiality of terms will function as a precondition for our meaning attitude: for any two supposedly referential terms e_1 and e_2, one cannot rationally assert that e_1 means e_2 without also believing that e_1 and e_2 are co-referential.

The attitudinal view of meanings also helps us to deal with the problem of non-trivially true identity statements: it suffices to say that for a true identity statement of the form "e_1 is e_2" to be non-trivial, it is necessary and sufficient that the corresponding meaning assertion "e_1 means e_2" be false when judged on the present account of meaning. The falsity of "e_1 means e_2" should result from one of the following two possibilities: either the hypothetical synonymy of e_1 and e_2 has been considered and refuted, or it has not yet been considered. In either of these cases, we can understand the identity statement "e_1 is e_2" formed by the co-referential terms "e_1" and "e_2", and recognize it as true, only by appeal to referential relations. In fact, our recognition of the truth of the statement "e_1 is e_2" is a necessary precondition for accepting assertions of the form "e_1 means e_2". We can thus explain non-trivially true identity statements without needing to posit meanings as transcendent truthmakers (or falsitymakers) for them and their like, and also without recourse to any viciously circular reasoning.

Besides playing a role in possible solutions to the above two classical problems, another indispensable function of meaning assertions will be particularly important in terms of directing us towards the pursuit of an attitudinal view of meaning: namely, that of allowing us to make some argument schemes that are intuitively valid analytically so. Let e_1 and e_2 be two expressions in our language, and now consider the following two arguments:

(1a) e_1 has the same meaning as e_2.

(1b) *a* knows (or believes) that …e_1…

(1c) Therefore, *a* knows (or believes) that …e_2…

(2a) e_1 has the same meaning as e_2.

(2b) *a* acts on the hypothesis that …e_1…

(2c) Therefore, *a* acts on the hypothesis that …e_2…

where …e_1… and …e_2… are declarative sentences in our language, such that the second sentence is obtained from the first by replacing one or more occurrences of e_1 by e_2.

Neither (1c) nor (2c) follows necessarily from the given premises, as it is possible that *a* does not bear the appropriate sort of cognitive relation to the meanings of e_1 and e_2. On the other hand, consider the following two arguments:

(3a) *a* means e_2 by e_1.

(3b) *a* knows (or believes) that …e_1…

(3c) Therefore, *a* knows (or believes) that …e_2…

(4a) *a* means e_2 by e_1.

(4b) *a* acts on the hypothesis that …e_1…

(4c) Therefore, *a* acts on the hypothesis that …e_2…

Both consequences—i.e. (3c) and (4c)—now follow necessarily from their premises. Once we are able to interpret the premise "*a* means e_2 by e_1" as ascribing to *a* an appropriate attitude towards the expressions e_1 and e_2, the information given in the premises allows us to leave out the possibility that would invalidate (1) and (2).

Neither the idea of doing away with meanings as entities, nor the idea of focusing upon synonymy as the central notion of the theory of meaning, is new (see Quine, 1951, pp. 22–23). Having said that, Quine argued strongly against the view that synonymy can be explained in terms of interchangeability *salva veritate*, for the reason that the latter requires analyticity. Nevertheless, the present view, while making synonymies the basis for a theory of meaning, escapes Quine's criticism. It seems that his criticism counts powerfully against views that focus on exact synonymies while also seeking to establish the interchangeability *salva veritate* of terms by deriving this from their extensional agreement. On the other hand, synonymies are now to be regarded not as necessary conclusions solely to be derived from the extensional agreement or prior usage of terms, but rather as hypotheses entertained on the basis of both prior facts about the relevant terms and the anticipation of success. Therefore, hypothesizing synonymies cannot be considered a free-floating language game. This approach to meaning can be properly characterized as pragmatist, insofar as it satisfies the three Deweyan desiderata—namely accessibility, hypotheticality and progressivity—that, as we saw earlier, together count as a clear expression of the principal methodological tenets of pragmatism. As a result of its compliance with the progressivity principle, the present view prompts us to conceive of the phenomenon of semantic change as a philosophical problem, and to regard theorizing the guiding principles for constructing better webs of meaning for our languages as the main challenge facing a philosophical approach to semantics.

2. A Theory of Meaning as a Theory of Justification

Once we have discovered that our planet is the third one from the sun, we may consider pairing the expressions "Earth" and "the third planet from the sun" as synonymous. Our decision should be based on a comparison of the anticipated consequences of our current alternatives. It is important to take into account the fact that we need not synonymize every expression of our language with some other expression. Indeed, many expressions can function without entering into synonymy relations with others. Many, including those that we understand through paradigmatic examples or partial explanations, belong to this category. The possibility of leaving some particular putative case or other of synonymy undecided matters, because any meaning attitude will bring with it a requirement of interchangeability for the relevant paired expressions in almost all contexts. Once we let the expressions "Earth" and "the third planet from the sun" be synonymous, we are obliged to extend our present set of beliefs with the addition of many others that will be obtained from our current beliefs just by exchanging "Earth" for "the third planet from the sun" and vice versa.

The Archimedean point from which one should proceed when seeking to develop further the attitudinal approach to meaning is this: that we regard a natural language as a critical tool useful for a wider network of activities, and treat meaning choices as improvements to the language in question. From this point of view, constructing, using and revising a natural language are goal-directed activities, so that Hintikka's distinction between definitory and strategic rules (see Hintikka, 1989, §3) will apply. The following passage from Hintikka and Sandu presents this distinction thus:

> In practically all such activities a distinction can be made between two different kinds of rules. This distinction is especially clear in the case of games of strategy […]. In them, we can distinguish the definitory rules which specify what may happen in the game from the strategic rules which tell how to play the game better or worse. For instance, the definitory rules of chess determine what moves are possible, what counts as checking and checkmating, etc., whereas following the strategic rules of chess is what makes a player better or worse. Strategic rules are not merely heuristic. They can in principle be as precise as the definitory rules, even though they are quite often so complicated as to be impossible to formulate explicitly. (2007, p. 20)

An example given by Hintikka offers further guidance when it comes to correctly evaluating the status of our meaning rules. As Hintikka and Sandu (2007, p. 20) note, in logic, inference rules should be counted as definitory rules: their function is to allow us to derive propositions without committing any fallacy. In a goal-directed activity, not all legal rules (i.e., those in compliance with the definitory rules) are effective with regard to moving us forward to reach the desired goal(s). Strategic rules then determine which legal moves are the right ones, and mastering these makes one a better player. Those possessing even just

a brief acquaintance with logic will know that starting with some arbitrarily chosen premises and applying inference rules in an arbitrary manner is not usually sufficient to derive a desired proposition, even if one applies the inference rules correctly. One should also master the strategic rules, which tell you which assumptions and inference rules to choose and in which order the inference rules should be applied.

If we consider natural language in the light of the above distinction, we can see that a general rule of substitution is an example of a definitory rule, as it is merely permissive: it lets us obtain new expressions from those previously given (on condition that we make no category mistakes). For example, from the proposition

(5) A square has four sides,

we may, by means of a substitution for the subject term, obtain the sentence

(6) A triangle has four sides,

and by means of a substitution for the predicate term we may produce the sentence

(7) A square has three sides.

Thus, restricted to declarative sentences and considered as an inference rule, substitution is not a sound rule: it does *not* guarantee that once we have accepted the original proposition, we shall obtain an acceptable proposition after performing a substitution. One may wonder why we have rules of this generality. In our example, one of the reasons is that such rules let us form as many expressions as we can from a single instance, and that ability is important for resource-bounded beings who are in need of languages of sufficient complexity, and who are expected to learn them. Moreover, we have a need to utter not just declarative sentences that we are ready to accept, but also those we are not prepared to accept.

What we have witnessed, then, is an example of a definitory rule for natural language: that of unrestricted substitution. Since the use of a language is a goal-directed activity, there should also be strategic rules telling us how to use our language effectively (relative to our goals). It should be easy to see that in the sense ascribed to them here, meaning rules are strategic rules: whilst unrestricted substitution allows us to construct as many expressions as possible from a given expression, meaning rules direct us to perform substitution in ways that are such that our language will better serve our goals. Considering the examples given above, while our general rules of substitution permit us to obtain both (6) and (7) from (5), both moves are bad from the perspective of our goal of arriving at knowledge—or, at least, that of extending our explicit true beliefs. On the other

hand, a rule which, for example, tells us to substitute "regular quadrilateral" for "square" compels us to make a good move in almost all cases.

Clearly one may, for instance, believe that "A square has four sides" without believing that "A regular quadrilateral has four sides", or vice versa. There could be someone, say *s*, such that for them the truth values of (8) and (9) below would differ:

(8) *s* believes that a square has four sides.

(9) *s* believes that a regular quadrilateral has four sides.

Moreover, even if an expression other than "regular quadrilateral" were to be chosen, we would still be faced with a similar conclusion. Therefore, it seems that one may generalize from this particular case and say that no strategic rule is good enough to be a sound rule. Even so, failure of substitutivity in intensional contexts does not raise a problem for the present view, as what we should really be seeking are those pairs for which substitution succeeds in as many contexts as possible. The point is that in that case we would be introducing a good strategic rule. In fact, it is part of our solution to decide whether we will consider such inferences as that of (9) from (8) as good: if we have developed a shared meaning-attitude towards the pair of expressions "regular quadrilateral" and "square", this need not mean that everyone has in fact mastered a rule to the effect that these terms should be used interchangeably. Rather, it means that if a competent user of our language, say, believes that a square has four sides, then he or she also should believe that a regular quadrilateral has four sides. On this basis, one is entitled, as a piece of practical reasoning, to infer (9) from (8), as in this case *s* would be viewed as following a strategic rule even though the possibility exists that he or she actually was not. One may fail while acting on the basis of having argued from (8) to (9), but it is a fact that in many strategic games, following a strategic rule does not and need not guarantee a win at the end of *every* actual playing of the game. (Beginner's luck seems to illustrate this point well.)

Since meaning rules are strategic rules within a goal-directed activity, they are supposed to facilitate the achievement of our desired goals in those of our activities that essentially require the use of language, and it is in this sense that they require justification: whether our meaning rules significantly contribute to the accomplishment of our goals should be the ultimate basis for their assessment. This also holds for our choice of theoretical constraints in the form of higher-order rules such as compositionality or contextuality. Such principles should only be adopted if we can be sure that following them while introducing lower-order rules will generally produce a better functioning language than the one already to hand.

Furthermore, as a result of construing meaning rules as strategic rather than definitory, the present view has a better chance than its competitors of staving off the criticisms directed against rule-based accounts of meaning. (These criticisms are especially forceful where use-based theories of meaning are concerned; see,

e.g., Gluer & Pagin, 1998.) Definitory rules are those required to make an activity—whether it be goal-directed or not—possible. Thus, if meaning rules were definitory, this would mean that they are required just to make use of the language possible, and it is easy to see that an understanding of rules of meaning in such terms leads quickly to a circularity. On the other hand, according to the present view, meaning rules are required neither to construct a language, nor to use it. Once we separate our theory of meaning from our theory of reference (Quine, 1951, pp. 22–23), it becomes possible that the role of meaning will be regulative rather than constitutive.

3. FROM EPISTEMIC TO EPISTEMO-SEMANTIC COHERENTISM

So far I have been arguing for the view that meanings should be construed as attitudes towards pairs of linguistic expressions, adopted by us on the basis of an anticipation of a better language, and that, as such, they should be considered amenable to justificatory evaluation. What remains is to elucidate the nature of the justification required for adopting particular choices of meaning attitude. In this section, after observing that a distinction that runs parallel to the foundationalism-coherentism distinction in epistemology is applicable to the justification of meaning attitudes, I will argue for the claim that a coherentist approach fits well with the present conception of meanings. Subsequently, I shall briefly outline how such an approach could be developed as an extension of epistemic coherentism.

As is well known, the two major approaches to epistemic justification—namely, epistemic foundationalism and epistemic coherentism—arise as alternative solutions to the regress problem: if every belief requires justification and a belief can be justified only by inferring it from some previously justified beliefs, then we should face the threat of infinite regress or circularity. As a rescue strategy, foundationalist theories of epistemic justification have had recourse to foundational or basic beliefs. These supposedly foundational beliefs are not in need of inferential justification, yet they can serve as premises in an attempt to justify a non-foundational belief. Epistemic foundationalism is subject to several strong criticisms, due to problems relating to the possibility of non-inferential justification for the foundational beliefs themselves, as well as the putative utility of such beliefs when it comes to imparting justification to others. Claiming that the foundational beliefs can be justified by an appeal to sensory experience, the foundationalist should be ready to engage head-on with the following Sellarsian dilemma: if our experiential states are non-doxastic, they themselves cannot justify any belief, while on the other hand, if they are belief-like and so are able to support other beliefs, they themselves require further justification. Either way, we fail to stop the justificatory regress (see BonJour, 1985, §6.2).

For many contemporary epistemologists who have found epistemic foundationalism to be untenable, epistemic coherentism offers a viable solution to the regress problem. The essence of the epistemic coherentist strategy is to bring into

play conditions applicable to systems of belief, so that a belief is justified insofar as that belief is internal to a belief system that satisfies these conditions. For the purpose of presenting the idea of epistemo-semantic coherentism, I will be content with the following most common list of epistemic coherence conditions:

a) The set of beliefs should be logically consistent.

b) The set of beliefs should have sufficient explanatory power.

c) There should be mutual inductive support among beliefs that are members of the set of beliefs.

Given that the coherence conditions are usually defined in terms of the internal properties of belief systems, coherentism is generally criticized for allowing the formation of belief systems in a void. Moreover, since it is most likely that we can form multiple coherent belief systems, there is a problem of criteria: how are we to choose among coherent systems without falling into some form of foundationalism? Alternative coherent systems of beliefs could even be jointly inconsistent. In other words, even if none of these alternative systems by itself implies a contradiction, they may jointly do so. Since no contradiction can be reasonably held, it is claimed that at least one of these allegedly coherent systems cannot be a faithful representation of reality. These arguments against epistemic coherentism owe their persuasiveness to the thought that it makes sense to posit a reality and truth that transcend human inquiry. Accordingly, the problems that these arguments are meant to imply are dissolved once we follow the Peircean strategy of redefining the relevant notions on the basis of intersubjectivity.

It is illuminating to consider an analogous problem of definitional regress: if giving an explicit definition of an expression by means of other (previously defined) expressions is the only way to determine the meaning of that expression in a satisfactory manner, then either an infinite regress or a circularity would be a threat once more. The oldest tradition, which clearly resembles the foundationalist solution for the epistemic regress, suggests that to solve the definitional regress we should accept some terms as semantically basic (or primitive).

To the extent that they make room for a realm of mind-independent entities in their metaphysics, most, if not all, foundationalist views of meaning typically accept some semantically basic terms on the basis of non-linguistic distinctions, given their conviction that the meanings of these expressions capture more basic ingredients of the mind or mind-independent reality itself. The traditional foundationalist strategy—represented, for example, by Aristotle—is to posit basic terms which can be given non-stipulative definitions that will be graspable without recourse to any further practices of explication (see Charles, 2000, sect. 10.6). In Aristotelian metaphysics, these basic terms correspond to basic entities with no proper parts, so they should only receive the simplest possible sort of definition (see Modrak, 2001, chap. 5). Locke's conceptualism echoes that view within epistemology, the main feature of the position (and those descended from it)

being that they base semantic distinctions on a distinction between simple and complex ideas. Russell's distinction between knowledge by acquaintance and knowledge by description provides another basis for distinguishing some expressions as basic. In each case, the resulting view would be vulnerable to criticisms similar to those directed against epistemic foundationalism.

There seems to be no explicitly foundationalist general theory of meaning that does not explain semantic basicness on the basis of a non-linguistic distinction. The most plausible explanation for this fact is that the views that seek to explain meanings of expressions on the basis of linguistic concepts are typically holistic ones. However, I wish to discuss a view of meaning—widely discussed in the philosophy of mathematics and the general philosophy of science—that represents a *prima facie* possible alternative foundationalist position where general theories of linguistic meaning are concerned. This more recent foundationalist view is based on the idea of definability by means of axioms.[4] Apparently, this view has the advantage that it does not depend on extralinguistic distinctions, while claiming that some terms are semantically basic. Consider, for example, the notions of point, line and plane in geometry. It is said that the role of axioms of a system of geometry is to determine the basic properties of these notions and thereby limit what propositions can be derived within the system (see Kline, 1980, p. 191). Generalizing from this idea, it is sometimes claimed that the axioms taken as a whole determine the meaning of basic terms by providing implicit definitions for them. Other notions can then be defined explicitly in terms of these notions. This option fails, however, as a result of Beth's theorem, which asserts the explicit definability of a predicate *F* relative to a theory *T* whenever *F* is said to be implicitly defined relative to *T*. (Beth presented this result in [1953]; see also the work of van Fraassen [2011] for further discussion.) Therefore, the implicit definability view is actually just a roundabout route to a foundationalist position.

Given that semantic foundationalism (in the above sense) is vulnerable to criticisms that are similar to those directed against epistemic foundationalism, and given that meanings can be considered genuine attitudes along with beliefs, I suggest that we consider developing a coherentist approach to the justification of meanings as a way to solve the problem of semantic regress. Applying the idea of epistemic coherentism to meaning attitudes, all of them are treated as equal members of the system, and only entire systems of meanings count as bearers of justification. Justification of a meaning attitude towards a pair of expressions is then only possible through a justification of the whole system of beliefs and meanings, including that meaning attitude itself. As in the case of

[4] As a result of the importance given to this idea in the writings of both Russell and Hilbert, this view is widely held in the philosophy of mathematics and logic. Quine attributes it to Gergonne, and notes that "it was still vigorous thirty years ago" (Quine, 1964, p. 71). See the work of Horwich (1997, p. 423) for an exposition and discussion of "the strategy of implicit definition".

epistemic coherentism, this holistic conception of the justification of meanings does not allow the definitional regress to get off the ground.

Accepting the idea that both beliefs and meanings are attitudes (the former towards propositions, the latter towards pairs of expressions), we may express beliefs and meanings with structurally similar statements. This fact allows us to present epistemo-semantic coherentism as a more or less straightforward extension of epistemic coherentism. For the sake of brevity, I shall make use of some formalism here. Instead of quoting expressions to form their names, I shall use for this purpose the schematic letters *e*, *e′* and *e″*. For any group *G* and for any member *a* in *G*, let $B_G p$ and $B_a p$, respectively, symbolize the propositions "*G* believes that *p*" and "*a* believes that *p*". The set of propositions $B^G = \{p: B_G p\}$ will be called the belief-set—or, for short, the *B*-set—of *G*. For the sake of presenting the basic idea, I shall content myself with Quinton's summative account of group knowledge (see Quinton, 1975), so that $B_G p$ if and only if $B_a p$ for all or most of the agents in *G*. For applications of the theory to various philosophical problems, one may, of course, work with more sophisticated notions of group belief. As a minimal rationality condition on belief, it should be assumed that belief sets contain basic tautologies, and that they are closed under some basic logical inference rules. This assumption can be justified by means of the second coherence condition: indeed, it is impossible for a set of beliefs to have sufficient explanatory power unless it incorporates a significant part of logical reasoning.

As a step in the direction of a formulation of epistemo-semantic coherentism, we may first extend our formalism so as to cover meanings in addition to beliefs. Let *G* be a group of users of a language (considered as a single entity) that, in addition to forming beliefs about propositions, can develop shared meaning attitudes towards pairs of expressions of their language. Let $\mu_G(e, e')$ denote the assertion that *G* means *e′* by *e*, where *e* and *e′* are *names* of two expressions of the language. Therefore, the set of pairs

$$M^G = \{(e, e'): \mu_G(e, e')\}$$

will be a binary relation on the set of expressions of the language. We shall call it the meaning set—or, for short, the *M-set*—of *G*.

Combining the belief-set and the meaning-set of a group, we obtain the belief-and-meaning set, or *BM*-set, of group *G* as the set $S^G = B^G \cup M^G$. We would obviously want our *BM*-set to be a harmonious whole:

Definition 1. We shall say that an expression is *substitutable* in a proposition *p* if it occurs as a stand-alone expression in *p* (i.e. it does not occur as a part of an idiomatic expression, a quotation, or a composite technical term) and it is only used (not mentioned) in *p*, except possibly in meaning-contexts such as the sentence "*e* means *e′*". We shall say that a *BM*-set S^G is substitution-consistent if, for every proposition *p* and every substitutable expression *e* in *p*,

$$\mu_G(e, e') \rightarrow (B_G p(e) \leftrightarrow B_G p(e')), \qquad (3.1)$$

or, equivalently,

$$(e, e') \in M^G \rightarrow (p(e) \in B^G \leftrightarrow p(e') \in B^G). \qquad (3.2)$$

Therefore, substitutional consistency amounts to a closure condition on the belief set B^G: to be substitution-consistent, B^G should include every substitution instance $p(e')$ for every proposition $p(e) \in B^G$ for all or most $a \in G$, given that G means e' by e.

The notion of coherence can now be extended to cover meaning, by means of the following auxiliary notion of substitutional consistency:

Definition 2. A system $S^G = B^G \cup M^G$ of beliefs and meanings of a group will be coherent if

a) B^G is epistemically coherent and

b) S^G is substitution-consistent with respect to M^G.

Finally, I suggest the following as a minimal definition of the coherentist justification of meaning:

Definition 3. A group G will be justified in meaning e' by e if and only if $S^G = B^G \cup M^G$ (where $(e, e') \in M^G$) is a coherent *BM*-set.

The notions above constitute the basic view that can be called epistemo-semantic coherentism. We may strengthen this basic view by adopting additional conditions on belief, meaning, and relations between belief and meaning. Developing the position in the direction of the first possibility is something that will concern the field of epistemic logic. To give an idea of what it would mean to pursue the second direction, let us consider how one might find support for the intuition that meaning relations should be equivalence relations: that is, for every group G, the relation μ_G that represents the totality of their meaning attitudes should satisfy

a) For every expression e, $\mu_G(e, e)$ (Reflexivity)

b) If $\mu_G(e, e')$, then $\mu_G(e', e)$ (Symmetry)

c) If $\mu_G(e, e')$ and $\mu_G(e, e'')$, then $\mu_G(e, e'')$ (Transitivity)

One may argue in favour of any of these conditions by showing that if we revise a system in accordance with that condition, then the resulting system is at least as coherent as the initial one. Thus, let $S^G = B^G \cup M^G$ be a coherent system.

This means that it is substitution-consistent, and B^G is epistemically coherent to a significant degree. To make the meaning relation reflexive, we let the new meaning relation be the set $M^G \cup \{(e, e)\}$. Since the addition of a pair (e, e) to M^G—even if it is not already in it—does not result in the addition of a new substitution instance for a group belief, the new system is also substitution consistent. Thus, adding reflexivity leaves B^G and the degree of epistemic coherence of the system as before. It follows that the resulting *BM*-system possesses the same degree of coherence.

Now let $S^G = B^G \cup M^G$ be a coherent system, and let $\mu_G(e, e')$. If e is substitutable in p(e), so is e' in $p(e')$. Moreover, given our definition of substitution-consistency, the assumption that $\mu_G(e, e')$ does not bring into play any new group belief. It follows that symmetry of meaning is an acceptable property. Acceptability of transitivity as a property for meaning relations can also be easily established. (Note that the summative account of group knowledge requires that the union of two insignificant groups of sceptics still be insignificant with respect to the beliefs of the group.)

I conclude this section with a remark on compositionality. Though I doubt whether meanings for natural language allow for any straightforward form of compositionality, it seems that a version of the compositionality principle can be consistently incorporated into our belief and meaning systems. In this case, it would read as follows: let S be a sentence with an immediately substitutable constituent c; then, if $\mu_G(c, c')$, then $\mu_G(S, S(c'/c))$ (where $S(c'/c)$ denotes the sentence which obtains by replacing one or more occurrences of c by an occurrence of c').

4. STRATEGIC MEANING REVISION

A natural language is basically a tool used by a group living in a dynamic environment. Therefore, every language needs revising at some points, in accordance with changes in that environment. Indeed, this is what we may actually observe: we add new expressions to the language, or some expressions eventually become obsolete, or we decide to give a new sense to an expression already in use with some other sense(s). In historical linguistics, these and similar phenomena are usually studied under the mantle of investigations into semantic change. In this section, I shall be considering semantic change from the coherentist point of view. In particular, I hope to demonstrate that:

a) Though semantic change is usually *stated* in terms of the meanings of linguistic expressions, the phenomena associated with semantic change can easily be reformulated and, more importantly, *explained* in terms of change in our meaning attitudes.

b) Reformulated in these terms, semantic change can be plausibly explained as a *purposive* process that results from the strategic cooperation of lan-

guage users aiming at maximizing the efficiency of their language through making better meaning choices. This should be the main difficulty both for transcendent views of meaning that work with unchanging entities, as well as views that work with innocuous entities, insofar as these entities change in ways that are incompatible with the phenomena associated with semantic change.

c) Owing to the fact that coherence allows of degrees, the coherentist view of meaning is able to explain why semantic change is and should be *continuous*; or, in other words, why meanings in natural language are and should be continuously changing.

Semantic change comes in many types, and there are varieties even within a single type. Below, I give some widely known examples of semantic change that, I believe, are sufficient to capture the general idea.

a) *Metonymy* is the addition of a new sense to an expression in such a way that though this new sense was originally not present, it is closely associated with the original meaning of that expression. This is obviously the case when, for example, "pen" is used in the sense of "writing", and "the sword" is used in the sense of "brute force", in the sentence "The pen is mightier than the sword".

b) *Widening* is defined as the process of extending the applicability of an existing expression: e.g., by extending its denotation or broadening its use to new contexts. For example, while "dog" originally meant a specific breed of dog, it later came to mean all breeds.

c) *Narrowing* is the opposite of widening. It happens either by restricting the denotation of a word to a part of its original denotation, or by restricting it to more specific contexts. For example, while "wife" originally meant "woman", "female", or "lady", its sense has been narrowed to "female spouse".

We can see that types of semantic change can be explicated as changes in respect of meaning attitudes:

a) Allowing a new expression to be linked with some expression e by extending the means relation: that is, we add the pair (e, e') and adopt a new attitude, which is that by saying e we mean e'. This will count as an appropriate action insofar as the expression e retains its previous senses.

b) Forgetting an existing expression linked with some expression e and thus restricting the means relation: that is, we erase the pair (e, e') and relinquish our attitude that by saying e we mean e'. We take this action if we are reluctant to retain the existing usage.

c) A combination of the above: forgetting an existing means attitude and adopting a new one in its place. This represents a change of our meaning attitude from "by saying *e* we mean *e′*" to "by saying *e* we mean *e″*".

Which action we should take depends on whether we want to keep on with the existing meanings or not. If the original relation is retained after the semantic change, the result will be what is known as polysemy.

After this brief explanation of the phenomena associated with semantic change within the attitudinal framework, I now wish to return to the second of the three statements put forward at the beginning of this section, where I claimed that semantic change can be plausibly explained as a *purposive* process that results from the strategic cooperation of language users who aim at maximizing the effectiveness of their language through making better meaning choices. I believe that this claim can be demonstrated within the present attitudinal-coherentist framework. The point is that if we should adopt a meaning attitude, it should be coherent with the rest of our beliefs and meanings. Moreover, the coherence conditions are defined in such a way that, as the degree of coherence of our system of beliefs and meanings increases, our language develops into a better instrument. Thus, though we may choose our *new* namings freely, we should be cooperating continuously in a way that will facilitate a positive evolutionary selection among meaning attitudes, where only the better ones survive. I now would like to consider two worries that can be raised in connection with this positive evolutionary stance towards semantic change, as discussing these will, I believe, lend more substance to the above outline of the notion of semantic change.

Firstly, one may question how it could be that such an extensive phenomenon as semantic change can be characterized in general terms as evolving in a positive sense. For there seem to be cases where it appears that we adopt a meaning choice that does not imply any increase in the effectiveness of our language. Establishing a slang term in order to discriminate against some group of people within the population seems to be an obvious case in point.[5] To deal with this problem, we must reflect on what it means to add a new word to our language. First of all, it does not by itself imply an immediate revision in respect of our meaning attitudes. Any claim to the contrary would most likely result from a confusion between naming and meaning. Quine alerted us to the fact that confusing naming with meaning may mislead one into hypothesizing that for a singular term to have the meaning it has, the object that is supposed to be named by it must exist (1948, pp. 28–29). At the same time, though, confusing naming with meaning may also mislead one in the other direction. This happens when one thinks that an expression *must* have a meaning if it is to successfully name some

[5] I am grateful to an anonymous referee of this journal for bringing this worry to my attention.

object. Here, the point is that we may enlarge our vocabulary without at the same time adopting additional meaning-attitudes in the sense of pairing new expressions with some that already belong to our previous vocabulary. Introducing a new word just to refer to an object will be less significant in comparison to adopting a meaning attitude for that word, and in the latter case it is considerably harder to anticipate how our language will work as a whole. In the case of introducing a new word by means of a conventional definition, however, both the reference(s) (extension) and the m e a n i n g (intension) of that word are established at once. When the word "circle" is introduced into the language by saying that "a circle" is/means "a set of all points in a plane with the same distance from a given point", its extension is fixed as the set of all sets of points satisfying this condition, and we declare that by saying "circle" we mean "a set of all points in a plane with the same distance from a given point". The case where a new word is established as the name of an entity is different, and if we are to respect the difference between naming and meaning, we should accept that when we make up an expression in order to label an entity, we do not automatically create a word endowed with a meaning. To see what happens in these cases, imagine that a group g_1 of scientists make up an expression e_1 to denote another group g, such that g is discriminated against by yet another group g_2 that introduces a word e_2 to refer to g. It is by way of the usage of e_1 within the context of the scientific enterprise of members of g_1 that e_1 becomes a technical term, and by way of the usage of e_2 within the context of the colloquial discourse of members of g_2 that e_2 attains the status of a slang term. (The possibility that these groups and their discourses may overlap does not seem to affect the cogency of the present argument.) Given that there are a significant number of people belonging to g_1 or g_2, we cannot accept it as a fact that $\mu_G(e_1, e_2)$ where G denotes our community itself (which is a larger group $G \supseteq g_2 \cup g_2$), as neither $\mu_{g1}(e_1, e_2)$ nor $\mu_{g2}(e_1, e_2)$ obtains. This shows that meaning involves more than naming: in order that a new meaning attitude can be established for two expressions, we must be able to anticipate that using them interchangeably will promote a more coherent system of beliefs and meanings.

Secondly, the fact that semantic change is usually a lengthy process with many people involved may cast a shadow over our claim as regards its purposiveness. Nevertheless, Davenport's (1960) game-theoretical analysis of Jamaican fishing communities has shown that there are cases where the actual strategy of a large group of real-life agents playing an iterated game-against-nature may accord with the technical game-theoretical analysis of the situation, and this supports the idea that the *prima facie* unorganized decision-making activity of a group may yet be purposive. Interpreting meaning decisions as strategic rules in games, semantic changes will correspond to changes in strategy. Indeed, a brief look at the history of strategic games suffices to lead to the conclusion that change in strategic rules happens quite often. Recall that strategic rules are those related to mastering a game with respect to the aims specific to that game; it is not knowledge of definitory rules, but following the appropriate strategic

rules, that makes you a good player. These rules develop through time: it was not until Murphy that the idea of d e v e l o p m e n t was known to chess players. So, at a time when strongly defensive or offensive moves were the ones considered good, if there were then to be an occasion to make such a move, making instead a move of the kind that we, now (after Murphy) would call a "developing move" might well have been deemed bad or ineffective.

In most cases, even the definitory rules of games are subject to change. (New rules may be added, some rules may be found to be redundant and dispensed with, or an existing rule may be altered.) However, this kind of change is relatively hard to meet with. When a definitory rule has been changed, we may decide whether we are now playing a new game or still playing the old game in newly evolved form. It seems to be down to us to decide. Many games have retained their names, even though their list of definitory rules has been revised considerably over the course of time. Many games with similar but different rules share the name "Chess", while some other variants, such as Chequers, Baroque and Take-All, which can all really be considered variants of chess, have acquired different names.

Why do we bother thinking of establishing a new rule, changing an existing rule, or even abolishing a rule in a game of strategy? Some reasons behind such changes in strategic games are internal: they refer only to other factors in the game, such as improving the applicability of the rules themselves. For example, when an arbiter was introduced into chess tournaments, this was because some other rules of the game could not be applied without an authority whose decisions during the course of playing the game would count as final. When the reason is external, it rather refers to whatever function that game has within some larger framework of human enterprise. For instance, the main reason for introducing a chess clock is to make the game more efficient and fun. Obviously, in neither of these cases was the new definition aimed at making the playing of the game somehow more in keeping with the essence of what it is to be a playing of the game of chess. This is not to deny, for example, that it is thanks to the definition of "arbiter" that we are willing to assent to the sentence "Lothar Schmid was the chief arbiter at the World Championship match between Bobby Fischer and Boris Spassky in Reykjavik in 1972".

That semantic change should be considered an ongoing purposive process realized by the cooperation of language users is something we have yet to clearly establish. In the first instance, this claim can rest on the analogy already drawn above between semantic change and change to the strategic rules of strategic games. To give an example, the main motivation behind the introduction of a new strategic rule in chess is surely to obtain a more effective winning strategy. Given that our use of language—at least by virtue of playing an essential role in our goal-oriented activities, such as producing, communicating and storing knowledge—is itself a goal-oriented activity, we may conclude that semantic change should also serve the purpose of achieving our goals. Observations in historical linguistics (see Meillet, 1905; Ullmann, 1957; 1962) also support the

conclusion that semantic change occurs for various reasons. Work in this field has led to a classification of such reasons into three categories: technical or linguistic ones, historical ones, and psychological ones. Considering these, it seems safe to conclude that semantic change is directed towards the achievement of some goal(s)—and this is why one should talk of "reasons for" rather than "causes of" semantic change.

It would be unrealistic to assume that we can always achieve these goals once and for all by making a single decision regarding some meaning attitude or other. Rather, we should be ready to revise the meaning relations that we have woven into language in the light of either some new goals or the possibility of better accomplishing our existing goals. This explains why the revision of our strategic rules is an ongoing process: our expectation is that our particular games will develop better if we consciously or unconsciously try to keep them attuned to the larger human enterprise.

The coherentist approach is best suited to explaining the continuity of semantic change, mainly because the notion of coherence allows for degrees: one system may be more or less coherent than another, and our aim is to make our systems ever more so. We may exploit this property to articulate a guiding principle for choosing an appropriate belief from a number of alternatives. Given a system S and a set of alternative propositions $\{p_1, p_2,\ldots\}$, let $\langle S+p_i\rangle$ for $i \in \{1,2,\ldots\}$ denote the list of all belief systems that we obtain as a result of adding the new proposition p_i and making the adjustments needed to our existing belief system to keep our set of beliefs coherent (e.g., by abandoning some of our beliefs, or replacing them with less contentious ones). Whenever we are able to make this calculation starting out with our present set of beliefs, we reach a new belief system S' which is at least as coherent as S. It is rational to add p_i to our beliefs if $\langle S+p_i\rangle$ contains one of the resulting belief sets with the highest degree of coherence. Iterating this process as we are faced with new alternative sets of propositions, we obtain a sequence of belief systems S, S', S'', …, such that each system should be at least as coherent as the preceding one. Applying the same idea to systems involving meanings along with beliefs, we may decide between alternative meaning attitudes.

Our coherentist attitudinal view of meaning may thus be regarded as satisfying the three Deweyan principles of accessibility, hypotheticality and progressivity mentioned earlier. This follows from the conception of belief and meaning attitudes underlying our view, and the conditions for their justification.

Firstly, while it is the case that meaning attitudes involve psychological states, the objects of meaning attitudes, their satisfaction and realization conditions, are all described in terms that are related to accessible observations, such as those concerning the use of linguistic expressions.

Secondly, to see that hypotheticality is also satisfied, note that when a linguistic community adopts some meaning attitude $\mu_G(e, e')$, this does not mean that they have *discovered* the truth that e means e' in their language. Neither does the theory assign meanings to the expressions e and e'—even after the system

that includes the attitude $\mu_G(e, e')$ has been justified. Rather, on the basis of an anticipation of there being favourable practical consequences, the community decides collectively to use these terms interchangeably, with the expectation that in this way their language will develop for the better. Given that the satisfaction of the latter involves indefinitely many future phenomena, adopting a meaning attitude involves the entertaining of a tentative, postulational hypothesis, and no definite number of positive observations can conclusively validate it.

Thirdly, it should be clear from their general characterization that meaning attitudes are revisable, and that they should be revised both in the light of our observations and in accordance with our expectations as regards better success in the future. This, it seems fair to say, suffices to show that on the account presented here, meanings will be linked to a domain of observation in a way that also promotes further observations.

ACKNOWLEDGEMENTS

I sincerely thank the editors of "Studia Semiotyczne" for their time and effort in handling the manuscript of the present article. I thank two anonymous reviewers for carefully reading the manuscript and commenting constructively on it, where this helped me to improve it considerably. I would additionally like to thank Prof. Lou Goble for commenting extensively on an earlier version of the text, and I would also like to thank him for making available to me the abstract of his PhD. dissertation. I also wish to express my gratitude to Dr. Nur Kurtoglu-Hooton and Tim Hooton for having carefully proofread the manuscript.

REFERENCES

Beth, E. W. (1953). On Padoa's Method in the Theory of Definition. *Indigationes Mathematicae*, *15*, 330–339.

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.

Brandom, R. B. (1994). *Making It Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge, Mass: Harvard University Press.

Charles, D. (2000). *Aristotle on Meaning and Essence*. Oxford: Clarendon Press.

Davenport, W. (1960). *Jamaican Fishing: A Game Theory Analysis*. New Haven, Connecticut: Human Relations Area Files Press.

Dewey, J., Bentley, A. F. (1945). A Terminology for Knowings and Knowns. *The Journal of Philosophy*, *42*(9), 225–247.

Fodor, J. A., Lepore, E. (2001). Brandom's Burdens. *Philosophy and Phenomenological Research*, *63*(2), 465–482.

Glüer, K., Pagin, P. (1998). Rules of Meaning and Practical Reasoning. *Synthese*, *117*(2), 207–227.

Goble, L. (1967). *A Coherence Theory of Meaning*. (Ph.D. dissertation, University of Pittsburgh, Pennsylvania).
Grice, P. (1957). Meaning. *The Philosophical Review*, *66*(3), 377–88.
Grice, P. (1975). Logic and Conversation. In: D. Davidson, G. Harman (Eds.), *The Logic of Grammar* (pp. 64–75). Encino, CA: Dickenson Publishing Company.
Grice, P. (1978). Further Notes on Logic and Conversation. In: P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 183–97). New York: Academic Press.
Hintikka, J. (1989). The Role of Logic in Argumentation. *Monist*, *72*(1), 3–11.
Hintikka, J., Sandu, G. (2007). What is Logic? In: D. Jacquette (Ed.), *Handbook of the Philosophy of Logic* (pp. 14–39). Amsterdam: Elsevier.
Horwich, P. (1997). Implicit Definition, Analytic Truth, and Apriori Knowledge. *Noûs*, *31*(4), 423–440.
Horwich, P. (1998). *Meaning*. Oxford: Clarendon Press.
Horwich, P. (2005). Reflections on Meaning. Oxford: Clarendon Press.
Kline, M. (1980). *Mathematics: The Loss of Certainty*. New York: Oxford University Press.
Lewis, D. (1969). *Convention. A Philosophical Study*. Harvard: Harvard University Press.
Macbeth, D. (2010). Inference, Meaning, and Truth in Brandom, Sellars, and Frege. In: B. Weiss, J. Wanderer (Eds.), *Reading Brandom: On Making It Explicit* (pp. 197–212). New York: Routledge.
Meillet, A. (1905). Comment les Mots Changent du Sens. *L'année Sociologique*, *10*, 1–38.
Misak, C. (2014). Language and Experience for Pragmatism. *European Journal of Pragmatism and American Philosophy*, *6*(2), 28–39.
Modrak, D. (2001). *Aristotle's Theory of Language and Meaning*. New York: Cambridge University Press.
Peregrin, J. (2009). Inferentialism and the Compositionality of Meaning. *International Review of Pragmatics*, *1*(1), 154–181.
Pietarinen, A.-V. (2006). *Signs of Logic: Peircean Themes on the Philosophy of Language, Games, and Communication*. Dordrecht: Springer.
Pietarinen, A.-V. (2014). Logical and Linguistic Games from Peirce to Grice to Hintikka. *Teorema: Revista Internacional de Filosofía*, *33*(2), 121–136.
Quine, W. V. O. (1948). On What There Is. *The Review of Metaphysics*, *2*(5), 21–38.
Quine, W. V. O. (1951). Two Dogmas of Empiricism. *The Philosophical Review*, *60*(1), 20–43.
Quine, W. V. O. (1964). Implicit Definition Sustained. *The Journal of Philosophy*, *61*(2), 71–74.
Quinton, A. (1975). The Presidential Address: Social Objects. *Proceedings of the Aristotelian Society*, *76*, 1–27+viii.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
Sellars, W. S. (1948). Realism and the New Way of Words. *Philosophy and Phenomenological Research*, *8*(4), 601–634.
Sellars, W. S. (1953). Inference and Meaning. *Mind*, *62*(247), 313–338.
Szubka, T. (2010). On the Very Idea of Brandom's Pragmatism. *Philosophia*, *40*(1), 165–174.
Ullmann, S. (1957). *The Principles of Semantics* (2nd ed.). Glasgow: Jackson. Oxford: Basil Blackwell.
Ullmann, S. (1962). *Semantics: An Introduction to the Science of Meaning*. Oxford: Blackwell.
Vanderveken, D. (2011). Formal Semantics for Propositional Attitudes. *Manuscrito*, *34*(1), 323–364.
van Fraassen, B. C. (2011). Logic and the Philosophy of Science. *Journal of the Indian Council of Philosophical Research*, *27*, 45–66.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 83–104
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.06
Article

PIERRE STEINER *

UNTANGLING THE KNOT OF INTENTIONALITY: BETWEEN DIRECTEDNESS, REFERENCE, AND CONTENT

SUMMARY: The notion of "intentionality" is much invoked in various foundational theories of meaning, being very often equated with "meaning", "content" and "reference". In this paper, I propose and develop a basic distinction between two concepts and, more fundamentally, properties of intentionality: intentionality-T (the fact that a state is directed to some object) and intentionality-C (the fact that a state is contentful). Representationalism is then defined as the position according to which intentionality-T can be reduced to intentionality-C, in the form of representational (i.e. contentful) states. Non-representationalism is rejecting this reduction, and argues that intentionality-T is more fundamental than intentionality-C. Non-representationalism allows for a new layered view of the relations between cognitive intentionality and linguistic intentionality; this view is presented at the end of the paper.

KEYWORDS: intentionality, content, mental representation, reference, representationalism, non-representationalism.

"Intentionality" is a much abused word, and it
means a variety of different things.
(Fred Dretske, 1994, p. 471)

INTRODUCTION. RECONSIDERING TWO PROBLEMS

Words and sentences of natural languages have meaning or semantic properties. Utterances consist in producing tokens of given sentences whose types be-

* University of Technology of Compiègne; Sorbonne University, COSTECH. E-mail: pierre.steiner@utc.fr. ORCID: 0000-0002-0041-5973.

long to a particular language. As is well known, the full meaning of an utterance goes beyond the conventional meaning of the uttered sentence: it is modulated by context-dependence, saliency effects, shared beliefs, and illocutionary force, and is more fundamentally related to what the speaker means when she is u s i n g that sentence. Addressees are able to grasp the meaning of utterances, going beyond the literal meaning of sentences. In order to do so, they display various cognitive abilities, including inference and pragmatic processes such as enrichment, loosening and transfer (Recanati, 2003). Mental states and processes are thus both resources and targets of many communicative processes.

According to the Gricean model (Grice, 1989), the meaning of utterances must be understood in terms of speaker's meaning, and speaker's meaning in terms of the intention of the speaker to induce a belief in the audience by an utterance, accompanied by the audience's recognition that the utterance was produced with that very same intention. Understanding the meaning of an utterance is thus a way of figuring out what the speaker's intentions are. Nevertheless, other mental states, such as beliefs and desires, may also be the targets of communication. Speakers might want to communicate thoughts to hearers. Beliefs and desires are not intentions, but they are, along with intentions, intentional, in the sense of being endowed with intentionality. They are contentful.

Another facet of communicative processes is the ability of agents to share attention and reference. It is the case when the speaker's labelling of some specific target is recognized as such by the hearer. What the speaker says succeeds in directing the hearer's attention to the intended referent. More basically, speakers and hearers might have their attention attracted to, for instance, the same perceptually salient features of a situation or object; this shared attention might also be a requisite for successful referring. Shared attention is a case of shared intentionality: both agents are directed to the same properties or objects.

I have deliberately used here the notion of intentionality. As François Recanati noted in 1998 (Recanati, 1998), a characteristic feature of recent work in the Gricean tradition has been the explicit employment of concepts from (and the intention to contribute to) cognitive science. The concept of intentionality is one of them. It now seems obvious that "in order to theorise expression meaning (word/sentence meaning), the basic intentionality of thought needs to be taken into account" (Haugh & Jaszczolt, 2012, pp. 111–112). Nevertheless, as the same authors note, the proliferation of the notion of intentionality can at times "create analytical confusion". Indeed, the term "intentionality" is often used for qualifying both linguistic entities (names, sentences) and mental states, but it remains a very puzzling property. It is often defined by a cluster of other properties. Amongst these properties we can find content, meaning, reference, and representation. For instance, in his seminal paper *The Intentionality All-Stars*, John Haugeland started from an apparent equivalence between intentionality, representation and content:

> Intentionality is hard to get a glove on. It is often glossed as that character of some things (items, events, states, …) that they are "of ", are "about" or "represent"

> others […]. In a different terminology, to have intentionality is to have (semantic) content. (Haugeland, 1990, p. 383)

Consider also the opening lines of the entry on "Intentionality" by Alex Byrne in a classic encyclopedia: "Some things are *about*, or are *directed on*, or *represent* other things" (Byrne, 2006, p. 405; author's emphasis).

Is there an unjustified equation or a basic confusion in the literature[1] between having intentionality, being contentful, and representing or being a representation?[2] We do not need to endorse such an austere interpretation. These quotations rather express the mundane fact that representation, intentionality and content are often seen as interrelated properties. But how can we untangle these relations between intentionality, content, and representation? This is our first problem.

Our second problem concerns the relations between the intentionality of language and the intentionality of mental states. Or, more exactly, it deals with a classic answer that has been proposed to this problem: the answer according to which mental intentionality is explanatorily and ontologically prior to linguistic intentionality. If one equates intentionality with representation, it is true that mental intentionality and linguistic intentionality are very similar in aspect. As John Searle writes: "Intentional states represent objects and states of affairs in the same sense of 'represent' that speech acts represent objects and states of affairs" (Searle, 1983, p. 4).

Nevertheless, the idea that there is an explanatory priority of mental intentionality on linguistic intentionality is a fundamental tenet in contemporary philosophy of language and philosophy of mind (Fodor, 1975; Schiffer, 1972; Loar, 1981). According to this idea, the intentionality of uttered public language sentences ultimately derives from the contents of the beliefs, desires, … that they express. True, the mental intentionality of the speaker does not fix the meaning of the sentences she uses (this meaning is a matter of linguistic conventions); but it fixes how the utterance of a sentence has to be understood (Montminy, 2010, p. 2911). This idea encompasses the Gricean picture of the communicative process mentioned in the first lines of the paper.

Consider John Searle again: "Words in the sentences of the language have a form of intentionality that is itself derived from the intrinsic or observer-independent intentionality of human agents" (2001, p. 53).

As Searle says, the mind "imposes" intentionality on linguistic expressions: "I impose Intentionality on my utterances by intentionally conferring on them certain conditions of satisfaction which are the conditions of satisfaction of certain psychological states" (1983, p. 28).

---

[1] For other instances of an apparent conflation between intentionality, representation and/or content, see (Crane, 2003, p. 30; Burge, 1979; Searle, 1983, p. 5).

[2] "Representation" is also, in itself, an ambiguous term: it can mean a relation—the representation relation, and an entity—as when one speaks of mental representations. More on this later (note 3).

If one substitutes "meaning" or "representation" for "intentionality", it is even more difficult to ignore the importance of approaches according to which, for instance, linguistic meaning must be explained in terms of mental meaning, or the semantics of external representations must be derived from the semantics of mental representations.

This popular answer to the problem of the relations between mental intentionality and linguistic intentionality generates what Jeffrey Speaks (2006) called the "mentalist picture of intentionality" (MPI), according to which social facts about public language meaning are derived from facts about the thoughts of individuals, and these thoughts have intrinsic (i.e. observer-independent) intentionality.

My objective in this paper is to reconsider MPI in the light of the first problem presented above: many if not all versions of MPI rest on what I will call a representationalist model of intentionality (RMI), according to which mental intentionality supervenes on, or is equated with, mental representations or contentful intracranial structures. There are different ways of rejecting MPI: one can deny the existence of mental intentionality (Chomsky, 2000), but this is a very expensive and debatable solution. One can argue that the intentionality of mental states is constituted by the intentionality of language (Dummett, 1993), but this would entail that non-linguistic beings are unable to be intentionally directed to the environment: this is a costly solution too. An alternative view would be to argue that mental intentionality and linguistic intentionality are interdependent (Davidson, 1984). But as long as one endorses RMI, it seems difficult not to attribute to mental intentionality a foundational or guiding role with respect to language, in virtue of the fact mental intentionality is intrinsically contentful. But, as we will see below, there are good reasons not to endorse RMI: one can acknowledge there is a non-linguistic and basic form of intentionality, but that this basic intentionality is not a matter of mental representations; it is not naturally contentful. The possibility of interdependence between linguistic intentionality and mental intentionality can consequently be reconsidered: linguistic practices (and the intentionality they produce) require intentional creatures; but the intentionality of these creatures is not contentful or content-conferring in itself. On the contrary, contentful (mental) intentionality can only emerge in linguistic practices. This picture requires a distinction between two kinds of intentionality. This distinction already exists today, in an implicit form in emerging works in cognitive science. We need to make it explicit, by clarifying what "intentionality" can mean (this is a way of answering the first question presented above).

The structure of this paper goes as follows. In section I, I define a basic distinction between two kinds of intentionality, and clarify their respective relations with reference, representation and content. In section II, I present the representationalist model on intentionality (RMI) and arguments in its favor. I then present in section III very recent critiques of RMI. Their core is retained for introducing an alternative to MPI in section IV, and sketching a new kind of interdependence between mental intentionality and linguistic intentionality.

I. INTENTIONALITY, REPRESENTATION, CONTENT AND REFERENCE

Let us go back to the first question presented above: how can we untangle the relations between intentionality, content, and representation?

A first answer consists in saying that content and representation are required for intentionality to exist. A physical state can have intentionality—it is "about" or "of" something—only if it has content, and so is a representational state. This is the basic claim of RMI (Cummins, 1989; Morgan & Piccinini, 2018). But in the literature we can also find places in which intentionality can figure in the characterization of what it is for a state to be contentful or representational. For example, according to Tim Crane, "a representation (linguistic, pictorial or mental) is the representation it is partly because of what it is about" (Crane, 2001, p. 317). Intentionality can thus be an essential and individuating property of representations. For Georges Rey (2003, p. 106), it is even in virtue of intentionality that mental states and events can have the contents they do, and be about objects or states of affairs.

From these observations, one might believe that the literature on intentionality is built upon a deep but tacit divide between those who believe that intentionality is prior to content and representation, and those who believe that content and representation are prior to intentionality. But, here too, conceptual distinctions must be considered before the possibility of philosophical confrontations: it might rather be the case that there is *one* notion of intentionality that makes intentionality necessarily dependent on content or representation, and *another* notion of intentionality that makes intentionality only contingently related to content and representation.

Indeed, Jaegwon Kim (1996, p. 21) proposed a distinction between referential intentionality and content intentionality. Referential intentionality denotes the aboutness or reference of thoughts or linguistic states. Content intentionality concerns the fact some states have contents or meanings. Kim's distinction is precious, and deserves to be further developed. Nevertheless, a caveat must be made concerning the association between "reference" and "intentionality". As is well known, "reference" is a polysemic term. It can mean the *act* of referring ("what are you referring to?"; "what does 'Pegasus' refer to?"; Allan, 2010), but also a relation, the relation of reference. As a relation, reference is a real relation: it is grounded on the existence of both relata (Crane, 2013). Words and thoughts are real entities. Pegasus or the Fountain of Youth do not exist. Hence, "Pegasus" or a thought about the Fountain of Youth are words and thoughts that do not refer, or that fail to refer. Still, these terms are meaningful; they have sense. And they have aboutness; they are directed to some (non-existing) entities. Intentionality and reference are thus very distinct properties (Loar, 2003, pp. 253–254): there can be intentionality without reference. Thoughts that are intentionally directed toward an entity which is taken as existing will refer to the entity in question if and only if there is an actual entity that satisfies the presupposition of its existence (Horgan & Tienson, 2002, p. 528). In order to point to the aboutness or

directedness dimension of intentionality, it is wiser to use the term "target-intentionality" rather than "reference".

We can now rephrase Kim's distinction as a distinction between intentionality as being directed, pointing or targeting to some object (what I will call "intentionality-T"), and intentionality as having representational or contentful properties (what I will call here "intentionality-C"). Intentionality-T (object-directedness) and intentionality-C (representation, content) are not two different ways of describing the same property—namely intentionality—when it is instantiated by some states or events. They are two distinct—yet related—properties that may be instantiated (together or not) by the same state, an instantiation in virtue of which we imprecisely use the term "intentionality" for describing that state.

Intentionality-C is the fact an intentional state has content: it means something. But what is content? Classically, content is defined by truth conditions or satisfaction conditions—sometimes, and more minimally, by accuracy and veridicality conditions (Burge, 2010). It may also be identified with abstract semantic entities like meanings, Fregean senses, possible worlds, modes of presentations, intensions, or propositions. This content is what makes the state or event semantically evaluable.

An important debate consists in defining the sources of content: is representational content a natural phenomenon that can be exhibited by material states independently of an observer or of some inclusion in linguistic practices? Or is representational content necessarily related to the possession of linguistic concepts, or to the participation in linguistic and social practices? Some authors can claim that natural forms of intentionality-C are prior to the existence of linguistically articulated intentionality-C (this corresponds to the difference between intrinsic [natural] intentionality and derived [linguistic] intentionality [Searle, 1983; Fodor, 1987; Dretske, 1994]), while other philosophers will argue that natural processes such as co-variation, information or biological functions are not sufficient for providing intentionality-C (Hutto & Myin, 2015).

Intentionality-T is not an object (be it an existing entity, a fictional entity, or an intentional object): it is the fact a state is directed towards, aims at, or is about a specific object, property or states of affairs which is not a component of the state (even though you might need to mention it for describing and individuating the state). Metaphors such as "aiming at", "targeting" or "pointing" are supposed to suggest the core of intentionality-T, echoing the etymology of the word (cf. the latin verb *intendere*, "aiming at something"). Intentionality-T corresponds to the fact some states are outward-directed. Facts involving intentionality-T are relational facts, in the sense that they do not only concern one agent or one state. But a relational fact does not necessarily entail any real existence for relations as irreducible dyadic properties (Campbell, 1990, p. 97). As in the case of intentionality-C, there is a debate concerning the sources of intentionality-T: is intentionality-T a natural phenomenon that can be exhibited by material states independently of an observer or of some inclusion in linguistic practices? Is

consciousness fundamental for intentionality-T? Some authors can be eliminativist regarding the natural existence of both intentionality-C and intentionality-T (Rosenberg, 2013), whereas others can be eliminativist towards the natural existence of intentionality-C and conservative regarding the natural existence of intentionality-T (Hutto & Myin, 2013; 2017).

One can accept that—by definition—intentionality-C is at bottom contentful or representational, while intentionality-T does not equate with content and representation. Nevertheless, there are also many cases in which both intentionalities may overlap. The same state or event can exhibit both intentionalities. For instance, we typically say of perceptual states, belief states, maps and sentences that they are "directed to" or "refer to" some objects (things, propositions, situations, states of affairs) which may exist or not (hence truth or falsity), but also that they have representational properties or content (possibly conceived as intensions, modes of presentations, propositions or senses)[3] which prescribe how their objects are targeted. Linguistic states such as sentences necessarily have intentionality-T and intentionality-C—unlike some mental states as, for instance, emotions, which can be intentionally directed to objects without having content (Voltolini & Calabi, 2009, pp. 9–17). Nevertheless, the intentionality-T of a linguistic state is distinct from its intentionality-C, for two sentences can be about the same object (i.e. have the same extension) and yet exhibit distinct content (i.e. distinct intensions).

Once we have made this distinction between intentionality-C and intentionality-T, another basic question arises: are these two kinds of intentionality directly related?

II. THE REPRESENTATIONALIST MODEL OF INTENTIONALITY

How are intentionality-C and intentionality-T related, in terms of necessary and/or sufficient relations of requirement? The first answer we will consider is the following: intentionality-C is necessary for intentionality-T. There is no intentionality-T of a state without intentionality-C instantiated by that state. This answer exists in at least three different forms: indeed, the dependency may be factual (constrained by laws of nature), logical (proper to the nature of intention-

[3] May we say that the distinction between intentionality-T and intentionality-C corresponds to the distinction (mentioned in footnote 2) between representation as a relation (A represents B) and representation as an entity (A is a representation of B)? No, for the following reason: whereas it is part of the concept representation that a representation represents something, we can have cases where having intentionality-C and having intentionality-T are distinct properties: one can instantiate the latter but not the former (and conversely). Still, as we will see, the representational theory of intentionality will basically equate intentionality-T with "representing", and intentionality-C with "being a representation". But other theories will see the representation relation as derived from a basic kind of (non-representational) directedness.

ality-T) or metaphysical (involving the supervention of intentionality-T on intentionality-C).

Be it a matter of factual, necessary or essential dependence, why would one believe that intentionality-T is necessarily dependent on intentionality-C? Here is a reconstruction of what I consider to be the main answer of proponents of this approach.

Intentional states are sensitive to the modes under which their objects are presented: one can be intentionally directed to Venus as the Morning Star and not intentionally directed to Venus as the Evening Star even though "The Morning Star" and "The Evening Star" are different ways to describe the same entity. Let us give the name aspectuality to the fact every intentional state is a state which is directed on, or about something else under an aspect. Directedness is thus necessary but not sufficient for having intentionality-T: aspectuality is also required. Aspectuality denotes the fact that when an intentional state is directed towards objects, these objects are always presented (or targeted) from a certain perspective. A thing is what it is, regardless of the way it is seen, described, desired or conceived; while an intentional object as being what is thought, described, desired, perceived,… is always individuated from the perspective the agent (or the state, be it linguistic or mental) has on it, for instance under the form of a definite description.

For proponents of the necessary dependence of intentionality-T on intentionality-C, being directed to some object from a perspective requires—or is equated with—"having content" because it is the content of a state which specifies how this state conceives, apprehends or merely *presents* things, in various modes: conceptually, descriptively, propositionally, but also more minimally perceptually and spatially. For instance, according to Alva Noë (2005, p. 189), the fact perceptual experience is intentional amounts to the fact it "presents things as being such and such". This is equivalent, for Noë, to the fact that perceptual experience has content. For Noë and Thompson, the fact perceptual experience has intentional content is equivalent to the fact "it purports to represent the world *as being this way or that*" (2002, p. 11; my emphasis).

More broadly, for Tim Crane: "Every intentional state or episode has a content—the way it represents what it is about or directed on" (Crane, 2013, p. 4).

Under the form of predication (if it is propositional) or more generally of specification, content would provide the aspectuality which is essential to intentionality-T. The intentionality-T of a state is necessarily grounded on that state representing a thing as being a certain way, and thus on that state having representational content or intentionality-C.

RMI currently forms the most important and popular version of the claim that intentionality-T is necessarily dependent on intentionality-C (Cummins, 1989, chap. 1; Morgan & Piccinini, 2018). All versions of RMI consider that intentionality-T is necessarily dependent on intentionality-C, with three additional subtheses:

(a) In the case of non-mental states (sentences, maps, pictures, …), the intentionality-C and the intentionality-T of those states are derived from the intentionality-T and intentionality-C of mental states;

(b) In the case of mental states, the vehicles or material bearers of intentionality-T and intentionality-C are mental representations. For any state exhibiting intentionality-T, this state is a physically realized state that carries or bears content (intentionality-C); it represents things to be a certain way.

(c) Mental representations have their content (intentionality-C) and intentionality-T naturally: these properties do not require linguistic or social practices for existing.

Mental representations are defined as intracranial contentful structures: something is a mental representation if it is about something else as being a certain way. A physical state can have intentionality-T—it is "about" or "of" something—only if it has content, and so is a representational state. Pointing to or targeting an object or a state of affairs, for a state or organism, is necessarily representing it. In order for S to be intentionally related to O, there must be a mental representation of O in S (Field, 1978; Fodor, 1985). Being realist about intentionality-T requires being realist about the existence of mental content and mental representations (see for example Jacob, 1997, chap. 1).

There are, of course, different versions of RMI. Depending on the theoretical framework one considers, mental representations may be complete, inert, propositional, denotational, action and perception-neutral, stable, complex, detailed, discrete, amodal, syntactically structured or symbolic, or proper to a language of thought (Fodor, 1975). But they may also be built and used on the fly; they can be modal (even when they are categorical), minimal (content-sparse), partial, action-oriented, context-dependent, embodied, distributed, or sub-symbolic. Nevertheless, there are more important distinctions inside of RMI than distinctions pertaining to the format of the vehicles of mental representations. An important debate exists concerning the origins of the content (or intentionality-C) of mental representations: according to some theories, the origins of the content of mental representations are to be found in phenomenal consciousness (Kriegel, 2013). For conceptual role theories (Block, 1986; Harman, 1993), the content of mental representations (intentionality-C) is determined by the functional role occupied by these representations in one's cognitive economy. For tracking theories, the content of mental representations finds its origins in natural tracking relations existing between the representational vehicle and some worldly object or states of affairs: these natural relations can be causal relations, informational relations, resemblance relations, counterfactual dependence relations, or teleological relations. Versions of RMI may thus diverge on the definition of the origins of the content (or intentionality-C) of mental representations.

As said before, the idea that intentionality-T is necessarily dependent on intentionality-C may be declined in several forms, depending on the kind of de-

pendency one sees between intentionality-T and intentionality-C. Accordingly, one can expect that there can be different forms of RMI. For moderate representationalist theories, there is no intentionality-T without intentionality-C, but intentionality-C does not fix all the facts there are concerning intentionality-T. An example of a moderate view is content-externalism.

For content externalism, there are cases in which the fact a representational state is directed to an object is not determined by its intentionality-C. Content-externalism underlines the fact some kinds of representational states are directed to, or refer to objects in a way which is not determined by their contents. This is notably the case of indexical thoughts and sentences, or cognitive states about proper names, natural kinds,… What determines intentionality-T here is not the way the object is represented, it is the nature of the relations there are between the thinking/speaking agent and the worldly object she thinks/speaks about. These relations can be direct or causal relations, or indirect relations such as deference or testimony, and also be a matter of context-dependence. Content externalists can also claim that mental or linguistic contents may be individuated by properties of their objects: in this sense, the intentionality-T of a state can play a crucial role in the identification of the intentionality-C of the same state.[4] More radically, some content externalists can claim that singular thoughts built around demonstratives or proper names directly refer to their objects (Recanati, 1993),[5] up to the point that these objects can even be seen as constituents of the thought. Reference to these singular objects is not mediated by descriptions in virtue of which one attributes general properties to these objects. Nevertheless, no semantic externalist would deny the claim that intentionality-C is necessary for intentionality-T to occur. She would just insist that intentionality-C is not sufficient for producing and individuating intentionality-T (indexicality and environmental dependence must be taken into account) and that intentionality-T can play a role in the individuation of content (the identity of intentionality-C is partially fixed by the environmental variables that are the objects of intentionality-T).

III. NON-REPRESENTATIONALIST APPROACHES ON INTENTIONALITY

In recent years, there have been a growing number of critiques of RMI (Steiner, 2014a). These critiques do not only reject the claim according to which intentionality-T supervenes on intentionality-C: they more fundamentally reject the claim that any form of intentionality-T necessarily involves intentionality-C, and

[4] I here leave aside the issue concerning the difference between broad content and narrow content.

[5] More precisely, Recanati (1993, p. 130) argues that we cannot think about objects without a conceptual mediation, but that in the case of *de re* concepts, the thought in which those concepts occur characterizes the referent itself independently of the fact the referent would satisfy—or not—the concept which is used to think of it.

thus the representationalist claim that any form of intentionality-T necessarily involves natural content and mental representations.

More positively and more precisely, these critiques of representationalist theories of intentionality endorse some of—if not all—the following claims:

(a) Intentionality-T and intentionality-C may be properties of mental states and of linguistic states, but intentionality-T is more primarily a property of deeds, actions and behaviour; it is a property in virtue of which mental states, linguistic states, deeds, actions and behaviour are directed to objects (events, properties, things, goals,…). Object directedness remains the defining core of intentionality, but it is not exclusive to mental states or linguistic states;

(b) Organisms can exhibit intentionality-T without harbouring states exhibiting intentionality-C. Mental states and bodily states can have intentionality-T without having intentionality-C.

(c) Even if there are local cases of intentionality-T supervening upon intentionality-C (for instance descriptive sentences), the basic case of intentionality is intentionality-T as proper to an embodied and engaged organism: it is from this intentionality that other forms of intentionality, such as contentful intentionality, are derived.[6] Intentionality-T is necessary for intentionality-C to occur, for the intentionality-C of a mental or linguistic state could not exist if the organism harbouring or producing that state did not exhibit intentionality-T. There is a primacy of intentionality-T over intentionality-C;

(d) The vehicles or material bearers of intentionality-T and intentionality-C may be mental representations; but there are also cases where the vehicles of intentionality-T and intentionality-C merely involve mental representations, or do not even require mental representations. For instance, behavioural states of an organism may exhibit intentionality-T, without being representational or contentful.

In this non-representationalist perspective, intentionality-C can also be named s e m a n t i c i n t e n t i o n a l i t y : it is first proper to linguistic or language-like states. Intentionality-T is a p r a g m a t i c i n t e n t i o n a l i t y , since it is primarily related to the way organisms act in their environment, or to the way their mental states have a functional role in relation with action. Nevertheless, it cannot be equated with a mere practical involvement with material things, for we would then lose the objectifying or object-directed central feature of intentionali-

[6] See for instance the work of Schlicht, proposing to "treat intentionality as a feature of whole embodied agents (paradigmatically organisms) who can be directed at objects and states of affairs in various ways, while representation should be regarded as a feature of mental states (and their respective vehicles or underlying mechanisms)" (2018, p. 1).

ty on which many authors insist (Menary, 2009, p. 36; Rowlands, 2010, p. 196; Thompson, 2007, pp. 22–27; Hutto & Myin, 2013; 2017). Indeed, object-directedness remains here the basic scheme justifying why non-representational intentionality-T *is* intentionality.[7] Nevertheless, relating to something as an object is not here necessarily supported by a contentful act of specification, predication, representation, or intellectual construction: it is fundamentally related to having a perspective on the world, in virtue of some activity, deeds, goals and purposes. Intentionality-T as object-directedness displayed by a conceptually articulated thought or a linguistic sentence, and intentionality-T as object-directedness displayed by an organism or an action (in the form of practical engagement for instance), are not distinct properties according to non-representationalism: they are one and the same property (object-directedness), having various modes and places of existence, including different relations with intentionality-C. According to this non-representationalist perspective, RMI unduly overgeneralizes to all intentional states what is only and originally proper to some of them, namely linguistic states: (necessarily) having content or meaning.

An important distinction inside of non-representational theories of intentionality concerns the acceptance or the denial of the existence and theoretical relevance of mental representations as naturally contentful states. This debate generates differences inside of claim (d). Some positions will deny that mental representations must be involved in—and be explanatorily relevant when accounting for—a n y cognitive process. But they will not deny the existence and explanatory relevance of contentful states of mind which are derived from socio-cultural practices (Hutto & Myin, 2013; 2017). Other, more moderate positions, will deny that mental representations must be involved in—and be explanatorily relevant when accounting for—some cognitive processes, but will simultaneously argue that they must be involved when explaining other complex, high-level or "representation-hungry" cognitive tasks. They will not entirely reject the relevance of the property of representation: they will only argue for its dispensability in some cases, including cases of intentionality-T (Rowlands, 2010). In any case, there can be creatures which are intentionally directed to the world without this intentionality being grounded on mental representations in the creature.

Let us consider some examples of non-representationalist theories of intentionality.

Hutto and Myin (2013; 2017) reject the existence of natural mental content, and thus of natural intentionality-C. Hutto and Myin claim that respectable naturalistic theories cannot accommodate naturally contentful cognitive states (also called "mental representations"), so these states should be theoretically eliminated. The mainspring of radical enactivism's attack on representationalism is its focus on the failures of the project of naturalizing mental content. Since their

[7] Since Brentano at least, object-directedness is the core property of intentionality (Brentano, 1874, Book II, ch. 1, par. 9).

first book, Hutto and Myin insist that neither informational theories nor teleosemantics are able to provide a satisfactory non-intentional explanation of the emergence of semantic properties: either they beg the question by already coming with intentional notions, or they merely deliver covariation and indication, which are not sufficient for giving semantic or representational content. Unable to be integrated in the naturalistic ontology it claims to be a part of, the representationalist programme would be "plagued with toxic debt, financed by loans it cannot pay back" (Hutto & Myin, 2013, p. 160). Since mental content has no place in a naturalistic ontology, there are good reasons to think it does not exist as an entity conveyed or produced by natural processes, including subpersonal and intracranial ones. In addition, non-representational means and models are already available and plausible for explaining basic cognitive phenomena.

For radical enactivism, basic cognition is a matter of embodied engagements responding to wordly offerings or information in the environment (Hutto & Myin, 2017, p. 130). These responses do not involve contents; but they must be explained, and in particular their c o n n e c t i n g properties with the world. Facing this challenge, the explanans of REC is intentionality or more precisely U r - i n t e n t i o n a l i t y , the "most primitive form of intentionality" (2017, p. 96)—in my terminology a version of intentionality-T, or a pragmatic intentionality. Non-representational intentionality is for REC the basic operator that will ground an embodied, enactive and extended approach to cognition: "basic minds target, but do not contentfully represent, specific objects and states of affairs" (Hutto & Myin, 2017, p. 130). Ur-intentionality is targeted at objects, without meaning, saying or representing them. It is a property of "aiming at" or "pointing towards" worldly offerings (Hutto & Satne, 2015, p. 530, note 7). This intentionality is a real, natural (and naturalizable) and intrinsic property of organisms (and not of mental or physical states inside of these organisms). Organisms display Ur-intentionality independently of what one may think or say about them, and independently of their possible inclusion in socio-cultural practices. This intentionality has been shaped through ontogenetic and phylogenetic history (Hutto & Myin, 2013, p. 111; 2017, p. 108, 130). It is naturalizable from the resources of t e l e o - s e m i o t i c s . According to radical enactivism, states or organisms are targeted or directed at F's because such targeting contributed to the fitness of the organism's ancestor and is therefore the reason why the state or organism endures: the reference to biological functions is enough for naturalizing intentionality-T and defining differences between aligned and misaligned responses, or appropriate and inappropriate responses, but it does not bring intentionality-C, correctness conditions, representation or misrepresentation (Hutto & Myin, 2017, pp. 104–115).

I take Hutto and Myin's radical enactivism as being a clear (and thought-provoking) example of an important trend of many recent critiques of representationalism in the philosophy of cognitive science: criticizing the existence of mental representations (or natural forms of intentionality-C) is first and foremost a way to gesture at the existence and importance of a (non-representational, non-theoretical) variety of intentionality-T. Sometimes inspired by phenomenology

(Merleau-Ponty, Heidegger) or by Deweyan pragmatism, many authors now speak of "motor intentionality" (Dreyfus, 2002; Hudin, 2006), "somatic intentionality" (Sachs, 2014), "Ur-intentionality" (Hutto & Myin, 2013), "autopoietic intentionality" (Varela, 1992) "skilled intentionality" (Kiverstein & Rietveld, 2015), or "enactive intentionality" (Gallagher, 2017), different varieties of intentionality-T that are independent of, and prior to, conceptual, discursive or linguistic forms of intentionality-C. As seen above, there is an important displacement of intentionality-T: as object-directedness, it is now a property of embodied and engaged organisms, and not of mental or physical states inside of these organisms. In virtue of intentionality-T, organisms are smoothly coupled with their environment; behaviours, deeds, and gestures are directed towards the world, but without representing it or without passing by contentful thoughts or propositional contents (except when the coupling process meets important perturbations, and fosters the need for reflective cognition).

The topological distinction between intentionality-T and intentionality-C allows for a layered model of the mind. Intentionality-C exists at the level of linguistic states and linguistically contentful mental events, while intentionality-T is first and foremost a property at the level of behaviours (and only derivatively a property of linguistic states and mental events). Intentionality-C would be derived from this primary form of intentionality. Here is, for instance, Robert Brandom, talking about practical intentionality where we speak here of intentionality-T as a pragmatic intentionality, and about discursive intentionality where we speak here of intentionality-C as a semantic intentionality:

> We might distinguish between two grades of intentionality: practical and discursive. Practical intentionality is the kind of attunement to their environment that intelligent nonlinguistic animals display—the way they can practically take or treat things as prey or predator, food, sexual partner or rival and cope with them accordingly. Discursive intentionality is using concepts in judgment and intentional action, being able explicitly to take things to be thus-and-so, to entertain and evaluate propositions, formulate rules and principles. […] One might claim […] that discursive activity, from everyday thought to the cogitations of the theoretical physicist, is a species of practical intentionality (or a determination of that determinable), and indeed, one that is intelligible as having developed out of nondiscursive practical intentionality, while still maintaining that it is a wholly distinctive variety. (Brandom, 2011, p. 10)

Intentionality-T becomes the genus from which intentionality-C is just a kind. Of course, various questions arise: is there intentionality-C without language and culture, for instance? Should the road from intentionality-T to linguistic intentionality-C pass by intermediaries which are forms of non-linguistic intentionality-C? Classical challenges are also addressed to non-representational theories of intentionality, and especially the challenge of accounting for the aspectuality of intentionality-T without appealing to content, representation or intentionality-C (Steiner, 2019).

IV. Reconsidering the Relations Between Language, Mind and Intentionalities

As said before, from the perspective of non-representationalism, the confusion or conflation between intentionality, representation and content arises from the fact one overgeneralizes to all intentional states what is only and originally proper to some of them, namely linguistic states: having content or meaning. This conflation attributes content to all intentional states, and turns mental states into foundations of linguistic states: hence the mentalist picture of intentionality presented in the introduction. But this picture appears to be dispensable once we make a distinction between two kinds of intentionality, namely intentionality-C (or semantic intentionality) and intentionality-T (or pragmatic intentionality). Armed with this distinction, one may understand how language and mind are mutually interdependent. The interdependence between language and thought is ensured by the articulation between intentionality-T and intentionality-C.

Language, from an ontogenetic and phylogenetic point of view, does not come from nowhere (Tomasello, 2003; 2008). Linguistic intentionality - the fact sentences and utterances have intentionality-C and intentionality-T - requires intentional agents for being instituted and effective, but the intentionality-T of those agents does not require intentionality-C for existing and instituting the possibility of language. The intentionality-T of those agents is exercised in various cognitive skills and activities which are involved in the acquisition, transmission and use of language. Notably, these cognitive skills and activities display the object-directedness which is proper to intentionality-T, and without which it would be impossible for linguistic episodes (utterances, thoughts, speech acts, written symbols) to be referentially and pragmatically anchored in a shared environment:

- Perceiving events, properties and affordances of objects;
- Sensitivity to natural signs (pointing to events, associating distinct events);
- Sensitivity to gestures and facial expressions, and responses-detection;
- Shared and joint attention (as in pointing gestures);
- Imitation (imitating actions, but also norm-governed patterns of behaviour);
- Anticipation;
- Motor control;
- Coordination of action;
- Normativity, as exemplified in the ability to produce norm-governed behaviour, including normative use of artifacts.

These skills display intentionality-T, but not necessarily the sharing or attributions of intentions as distinct psychological states, in virtue of mindreading or a theory of mind. Some of these skills may be involved in communicational practices. But communication is not necessarily linguistic communication. And it is doubtful communication practices pass by the manufacture, the interpretation or the exchange of contents (and intentionality-C), especially in the form of naturally contentful mental representations. Those communicational practices allow for continuity between basic intentionality-T activities, and linguistically articulated contentful activities. As a way to coordinate and regulate action on the basis of signals, communication requires intentionality-T and the skills mentioned above, but it also establishes the possibility of social and cultural practices from which language (and intentionality-C) appears. It is at this linguistic level that reference, concepts, truth conditions, contextual detachability and objectivity emerge. In these previously mentioned basic intentionality-T skills, there are objects and objectivation, but not objectivity as a property of representations. Objectivity only appears with language, in relation with the ability to have thoughts having objective content, or a content independent of what we do when we think it (Davidson, 2001, chap. 9). There, it makes sense to speak about the world being contentfully presented to agents in certain ways, ways we can describe and predict with the use of concepts such as beliefs, intentions or thoughts (Steiner, 2014b).

There is another primary way by which language both arises from a framework of cognitive skills and reconfigures cognitive experience, not only adding one more function to this framework (speaking, talking, languaging). Since at least Vygotsky, much has been said about the effects of the acquisition of natural language for the development of new cognitive abilities, and for the reconfiguration of former abilities: reflexivity, memory, metapresentation, systematic reasoning, attention, and so on (Clark, 1998; Carruthers, 2002; Millikan, 2004, chap. 19). But one can also underline how much natural language (and especially linguistic concepts) is a condition for the development of intentionality-C and also new forms of intentionality-T.

Indeed, according to the non-representationalist perspective, intentionality-C is primarily a discursive intentionality: it is exhibited by linguistic states and by language users, and by agents that use linguistic representations for producing thoughts. It is only in virtue of linguistic competences and in virtue of the existence of linguistic practices that agents can have mental states exhibiting intentionality-C. But from this perspective, intentionality-C provides us with the possibility to be directed to new events and possibilities. In this sense, the possibility for an agent to produce some cognitive attitudes exhibiting intentionality-T may depend on mental intentionality-C, itself dependent on concepts and referential practices proper to language and its material inscriptions.[8]

[8] As Frege already remarked in 1882, “without symbols we would further hardly raise ourselves to the level of conceptual thought. In giving the same symbol to similar but

For instance, without the concept "electron", one cannot think about electrons, and entertain contentful thoughts about them. As Ruth Millikan claims, "merely having a word can be enough to have a thought of its referent" (Millikan, 2017, p. 35). Our ability to produce contentful thoughts exhibiting intentionality-T about electrons is enabled by our mastery of the concept "electron": what this concept means and refers to is not fixed by our thoughts, but by linguistic practices. Words express concepts which are public entities, shared by the members of a linguistic community. The referential anchoring of the concepts is ensured by the community (Kripke, 1980; Putnam, 1975; Burge, 1980), not by intrinsically contentful mental states of individuals. As David Kaplan remarked, words often come to us prepackaged with a semantic value. Typically, we are more consumers than creators of language and its intentionality. This is not a cognitive defect; on the contrary. For Kaplan, this fact allows us

> to broaden the realm of what can be expressed and to broaden the horizons of thought itself. On my view, our connection with a linguistic community in which names and other meaning-bearing elements are passed down to us enables us to entertain thoughts through the language that would not otherwise be accessible to us. Call this the Instrumental Thesis. […] It urges us to see language, and in particular semantics, as more autonomous, more independent of the thought of individual users, and to see our powers of apprehension as less autonomous and more dependent on our vocabulary. Contrary to Russell, I think we succeed in thinking about things in the world not only through the mental residue of that which we ourselves experience, but also vicariously, through the symbolic resources that come to us through our language. It is the latter vocabulary power that gives us our apprehensive advantage over the nonlinguistic animals. My dog, being color-blind, cannot entertain the thought that I am wearing a red shirt. But my color-blind colleague can entertain even the thought that Aristotle wore a red shirt. (Kaplan, 1989, p. 602)

Kaplan agrees that "to use language as language, to express something, requires an intentional act. But the intention that is required involves the typical consumer's attitude of compliance, not the producer's assertiveness" (1989, p. 602). By "compliance", Kaplan here notably refers to the important phenomenon of linguistic d e f e r e n c e, by which the division of linguistic labour allows one to defer to experts the ability to define the reference of some terms. However, I think we can go further than Kaplan in order to point to the foundational importance of another kind of (non-linguistic) intentionality for explaining language use. Not intentionality-C of course, for we have just seen that it is a by-product of language, just like some versions of intentionality-T which supervene on this intentionality-C. The intentionality that matters here is intentionality-T as

---

different things, we no longer symbolize the individual thing but rather that which they have in common—the concept—and the concept itself is first gained by our symbolizing it, for, since the concept is of itself imperceptible to the sense, it requires a perceptible representative in order to appear to us" (1882, p. 156).

a pragmatic intentionality enabling agents to be directed to others and to the environment, so that they can start mastering and sharing the contentful resources of natural language.

As we have seen, there are intentionality-T performances which are involved in the acquisition and use of natural language, including concept use. But there is no reason to think that these performances are grounded on the processing of mental contents or mental representations (Hutto & Myin, 2015). However, the fact these agents can entertain cognitive attitudes (intending, desiring, believing, attending…) towards distant, absent, counterfactual or abstract properties or states of affairs requires the use of concepts or public meanings, and is deployed by the production and the manipulation of linguistic and mental states endowed with intentionality-C. One may argue that utterances sometimes express or are driven by pre-existing propositional attitudes, but these propositional attitudes are first constituted by the use of natural language sentences, exhibiting intentionality-C and intentionality-T, and dependent on intentionality-T as pragmatic intentionality.

The capacity to have contentful thoughts, to be directed to distal or abstract events in virtue of contents, and the capacity to talk, arise and develop together, and they are dependent on intentionality-T as a pragmatic intentionality. There is no contentful thought (exhibiting intentionality-C) without natural language; but there is no mastery of natural language without intentionality-T in agents. The acquisition and use of natural language require intentional-T capacities in agents (joint attention, coordination, responsiveness, directedness, …) in a context of social and cultural norms, but it does not require states which would be intentionally contentful independently of language.

V. CONCLUSION

When John Perry writes that

> The intentionality of linguistic acts is a special case of the intentionality of purposeful action. The language to which a token belongs, the identity of the words and their meanings, the syntax, the reference of terms, all derive from the minds of the speakers, and *connections between those minds, other minds, things and properties*. (2006, p. 316; my emphasis)

There are at least two distinct ways of understanding what he means by "connections between those minds, other minds, things and properties"—or so I have claimed in this paper. Those connections can be intentional in two different ways: they can exhibit (semantic) intentionality-C which would determine intentionality-T, or (pragmatic) intentionality-T only. According to MPI and RMI, the intentionality of linguistic acts is dependent on individual mental representations. The intentionality-T of a cognitive state is a matter of intentionality-C. In this first sense, Perry's connections are contentful, and ground the possibility of linguistic intentionality. Contrary to RMI and thus MPI, and in continuity with

recent critiques of RMI in philosophy of cognitive science, I have argued that it is possible to hold simultaneously that public natural language is constitutive of one kind of intentionality (intentionality-C, and the intentionality-T which supervenes on it) and that another kind of intentionality (intentionality-T, a pragmatic intentionality) is necessary for the acquisition and use of public natural language. In this picture, the intentionality of linguistic acts is both constitutive and constituted: it is constitutive of thoughts and their intentionality-C, and constituted by intentionality-T as a property of deeds, non-linguistic acts and behaviour, historically and socially situated. One can appeal to a kind of non-linguistic intentionality for explaining linguistic intentionality without embracing mentalism and representationalism, but by developing a more pragmatic picture of the intertwinement between mind and language.

REFERENCES

Allan, K. (2010). Referring as a Pragmatic Act. *Journal of Pragmatics*, *42*(11), 2919–2931.

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

Brandom, R. (2011). *Perspectives on Pragmatism: Classical, Recent, and Contemporary*. Cambridge, Mass.: Harvard University Press.

Brentano, F. (1874). *Psychologie vom empirischen Standpunkte*. Leipzig: Duncker & Humblot.

Burge, T. (1979). Individualism and the Mental. In: P. A. French, T. E. Ueling, H. K. Wettstein (Eds.), *Midwest Studies in Philosophy,* (Vol. 4, pp. 73–122). Minneapolis: University of Minnesota Press.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Byrne, A. (2006). Intentionality. In: J. Pfeifer, S. Sarkar (Eds.), *The Philosophy of Science: An Encyclopedia*. Milton Park, Abingdon: Routledge.

Campbell, K. (1990). *Abstract Particulars*. Oxford: Blackwell.

Carruthers, P. (2002). The Cognitive Functions of Language. *Behavioral and Brain Sciences*, *25*(6): 657–719.

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Clark, A. (1998). Magic Words: How Language Augments Human Computation. In: P. Carruthers, J. Boucher (Eds.), *Language and Thought: Interdisciplinary Themes* (pp. 162–183). Cambridge: Cambridge University Press.

Crane, T. (2001). Intentional Objects. *Ratio*, *14*(4), 298–317.

Crane, T. (2003). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation* (2nd ed.). Milton Park, Abingdon: Routledge.

Crane, T. (2013). *The Objects of Thought*. Oxford: Oxford University Press.

Cummins, R. C. (1989). *Meaning and Mental Representation*. Cambridge (MA): MIT Press.
Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
Davidson, D. (2001). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
Dretske, F. (1994). If You Can't Make One, You Don't Know How It Works. *Midwest Studies in Philosophy*, *19*(1), 468–482.
Dreyfus, H. L. (2002). Intelligence Without Representation: Merleau-Ponty's Critique of Mental Representation. *Phenomenology and the Cognitive Sciences*, *1*(4), 367–83.
Dummett, M. (1993). *The Seas of Language*. Oxford: Oxford University Press.
Field, H. (1978). Mental Representation. *Erkenntnis*, *13*(1), 9–61.
Fodor, J. A. (1975). *The language of Thought*. New York: Crowell.
Fodor, J. A. (1985). Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum. *Mind*, *94*(373), 76–100.
Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge (MA): MIT Press.
Frege, G. (1882). On the Scientific Justification of a Concept-Script. *Mind*, *73*(290), 155–160.
Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press.
Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge (MA): Harvard University Press.
Harman, G. (1993). Meaning Holism Defended. *Grazer Philosophische Studien*, *46*(1), 163–71.
Haugeland, J. (1990). The Intentionality All-Stars. *Philosophical Perspectives*, *4*, 383–427.
Haugh, M., Jaszczolt, K. M. (2012). Speaker Intentions and Intentionality. In: K. Allan, K. Jaszczolt (Eds.), *Cambridge Handbook of Pragmatics* (pp. 87–112). Cambridge: Cambridge University Press.
Horgan, T., Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In: D. J. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings* (pp. 520–533). Oxford: Oxford University Press.
Hudin, J. (2006). Motor Intentionality and Its Primordiality. *Inquiry*, *49*(6), 573–590.
Hutto, D. D., Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge (MA): MIT Press.
Hutto, D. D., Myin, E. (2017). *Evolving Enactivism: Basic Minds Meet Content*. Cambridge (MA): MIT Press.
Hutto, D. D., Satne, G. (2015). The Natural Origins of Content. *Philosophia*, *43*(3), 521–536.

Jacob, P. (1997). *What Minds Can Do: Intentionality in a Non-Intentional World*. Cambridge: Cambridge University Press.
Kaplan, D. (1989). Afterthoughts. In: J. Almog, J. Perry, H. Wettstein (Eds.), *Themes from Kaplan* (pp. 565–614). Oxford: Oxford University Press.
Kim, J. (1996). *Philosophy of Mind*. Boulder, Colorado: Westview Press.
Kiverstein, J., Rietveld, E. (2015). The Primacy of Skilled Intentionality: On Hutto & Satne's the Natural Origins of Content. *Philosophia*, *43*(3), 701–721.
Kriegel, U. (2013). The Phenomenal Intentionality Research Program. In: U. Kriegel (Ed.), *Phenomenal Intentionality*. Oxford: Oxford University Press.
Kripke, S. (1980). Naming and Necessity. Cambridge (MA): Harvard University Press.
Loar, B. (1981). *Mind and Meaning*. Cambridge (MA): Cambridge University Press.
Loar, B. (2003). Phenomenal Intentionality as the Basis of Mental Content. In: M. Hahn, B. Ramberg (Eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge* (pp. 229–258). Cambridge (MA): MIT Press.
Menary, R. (2009). Intentionality, Cognitive Integration and the Continuity Thesis. *Topoi*, *28*(1), 31–43.
Millikan, R. (2004). *Varieties of Meaning*. Cambridge (MA): MIT Press.
Millikan, R. (2017). *Beyond Concepts. Unicepts, Language and Natural Information*. Oxford: Oxford University Press.
Montminy, M. (2010). Context and Communication: A Defense of Intentionalism. *Journal of Pragmatics*, *42*(11), 2910–2918.
Morgan, A., Piccinini, G. (2018). Towards a Cognitive Neuroscience of Intentionality. *Minds and Machines*, *28*(1), 119–139.
Noë, A. (2005). *Action in Perception*. Cambridge (MA): MIT Press.
Noë, A., Thompson, E. (2002). Introduction. In: A. Noë, E. Thompson (Eds.), *Vision and Mind: Selected Readings in the Philosophy of Perception*. Cambridge (MA): MIT Press.
Perry, J. (2006). Using Indexicals. In: M. Devitt, R. Hanley (Eds.), *The Blackwell Guide to the Philosophy of Language* (pp. 314–334). Oxford: Blackwell.
Putnam, H. (1975). The Meaning of "Meaning." In: H. Putnam (Ed.), *Mind, Language and Reality: Philosophical Papers* (Vol. 2, pp. 215–271). Cambridge: Cambridge University Press.
Recanati, F. (1993). *Direct Reference: From Language to Thought*. Oxford: Blackwell.
Recanati, F. (1998). Pragmatics. In: E. Craig (Ed.), *Routledge Encyclopedia of Philosophy* (Vol. 7, pp. 620–633). Milton Park, Abingdon: Routledge.
Recanati, F. (2003). *Literal Meaning*. Cambridge: Cambridge University Press.
Rey, G. (2003). Chomsky, Intentionality, and a CRTT. In: L. M. Antony (Ed.), *Chomsky and His Critics* (pp. 105–139). Malden: Blackwell.
Rosenberg, A. (2013). Disenchanted Naturalism. In: B. Bashour, H. Muller (Ed.), *Contemporary Philosophical Naturalism and its Implications* (pp. 13–17). Milton Park, Abingdon: Routledge.

Rowlands, M. (2010). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge (MA): MIT Press.
Sachs, C. B. (2014). Discursive and Somatic Intentionality: Merleau-Ponty Contra 'McDowell or Sellars'. *International Journal of Philosophical Studies*, *22*(2), 199–227.
Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.
Schlicht, T. (2018). Does Separating Intentionality from Mental Representation Imply Radical Enactivism? *Frontiers in Psychology*, *9*, 1–14.
Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
Searle, J. R. (2001). *Rationality in Action*. Cambridge (MA): MIT Press.
Speaks, J. (2006). Is Mental Content Prior to Linguistic Meaning? *Noûs*, *40*(3), 428–467.
Steiner, P. (2014a). Enacting Anti-Representationalism. the Scope and the Limits of Enactive Critiques of Representationalism. *Avant: Trends in Interdisciplinary Studies*, *5*(2), 43–86.
Steiner, P. (2014b). The Delocalized Mind. Judgments, Vehicles and Persons. *Phenomenology and the Cognitive Sciences*, *13*(3), 437–460.
Steiner, P. (2019). Content, Mental Representation and Intentionality: Challenging the Revolutionary Character of Radical Enactivism. *Croatian Journal of Philosophy*, *18*(55), 153–174.
Thompson, E. (2007). *Mind in Life*. Cambridge (MA): Harvard University Press.
Tomasello, M. (2003). *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge (MA): Harvard University Press.
Tomasello, M. (2008). *Origins of Human Communication*. Cambridge (MA)/London: MIT Press.
Varela, F. (1992). Autopoiesis and a Biology of Intentionality. In: B. McMullin (Ed.), *Proceedings of the Workshop 'Autopoiesis and Perception'* (pp. 4–14). Dublin: Dublin City University.
Voltolini, A., Calabi, C. (2009). *I problemi dell'intenzionalita*. Turin: Einaudi.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 105–134
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.07
Article

RICHMOND KWESI *

SEMANTIC MEANING AND CONTENT: THE INTRACTABILITY OF METAPHOR

SUMMARY: Davidson argues that metaphorical sentences express no propositional contents other than the explicit literal contents they express. He offers a causal account, on the one hand, as an explanation of the supposed additional content of a metaphor in terms of the effects metaphors have on hearers, and on the other hand, as a reason for the non-propositional nature of the "something more" that a metaphor is alleged to mean. Davidson's account is meant to restrict the semantic notions of meaning, content, and truth, to literal sentences. I argue that the Davidsonian causal account does not satisfactorily account for metaphor's figurativeness, speakers' assertion and hearers' uptake of metaphor, and our discursive practices of using metaphors in disagreements and argumentation. I offer a non-compositional analysis of a semantic account of metaphor within which one can make sense of the applicability of the notions of meaning and content to metaphor. This analysis shows that metaphorical sentences have meanings other than, and in addition to, their literal meanings and what speakers can use them to mean.

KEYWORDS: metaphor, compositionality, Davidson, meaning, content, causal account.

1. INTRODUCTION

Semanticists who have worked on the semantic notions of meaning, content, truth and assertion have had to grapple with the phenomenon of figurative language in general and metaphor in particular. Some of them have engaged with the following questions: Do metaphorical sentences have contents other than, or in addition to, the literal contents they express? If metaphorical sentences have

* University of Ghana, Department of Philosophy and Classics. E-mail: rkwesi@ug.edu.gh. ORCID: 0000-0001-5286-1935.

any such non-literal contents, are these contents propositional in nature? How do metaphorical sentences come to have such additional (propositional) contents? And, how do users of metaphors—speakers and hearers—associate, derive, or capture, these additional contents? An adequate semantic theory or a semantic account of our linguistic practices has to provide an answer to the above questions by either explaining away and dispensing with the issue of metaphor, that it is not amenable to the semantic notions properly understood, or that the semantic notions are not appropriately applicable to the phenomenon of metaphor, or that the senses in which the semantic notions are applicable to metaphor are different from those in which they are applicable to ordinary literal uses of language.

Let us consider an intuitive story of how the notions of meaning, content, and truth are indispensable to thinking about metaphor. In the sentence, "Gabriele is a crocodile—he is impulsive and angry, he changes like the wind", the first part of the sentence—"Gabriele is a crocodile"—is a metaphor;[1] and the second part—"he is impulsive and angry, he changes like the wind"—is considered variously as the content, meaning, interpretation, or paraphrase of the content of the metaphor. The metaphor is about a subject and a predication of a property to the subject, the predicate presents us with a description of the subject—it is an attribution of a particular property to the referent of the subject term of the metaphor. The metaphorical statement has a semantic value—it is true or false if the subject "fits" the description, or if it is the way in which it is being described. Taken literally then, the description is false (on the assumption that Gabriele is human and not a crocodile) but since the statement *is* a metaphor (or is being used as a metaphor), it is the second part of the remark—the paraphrase—which gives the interpretation of the property being attributed to Gabriele in literal terms that confers truth (or falsity) on the metaphor. That is, the metaphorical statement has truth value, and the truth value is derived from, and dependent upon, the truth or falsity of the corresponding interpretation or paraphrase of the metaphor. In this example then, the description of Gabriele as a crocodile is true if it is true that Gabriele is impulsive and angry and changes like the wind. This intuitive story is unpersuasive to Donald Davidson and other theorists who are sympathetic to his account of metaphorical meaning and content.

I discuss in this paper Davidson's treatment of the meaning and content of metaphorical sentences. Davidson maintains that metaphorical sentences express no propositional contents other than the explicit literal contents they express. He offers a causal account, on the one hand, as an explanation of the supposed additional content of a metaphor in terms of the effects metaphors have on hearers, and on the other hand, as a reason for the non-propositional nature of the "something more" that a metaphor is alleged to mean. In the analysis of the Davidsonian position, I will argue that what metaphors cause us to do, and the effects they

[1] Another way of talking about the sentence is to say that it is being used *as* a metaphor. The sentence itself could also be another figure of speech like irony or overstatement, or simply an insult, although both the linguistic and the non-linguistic contexts will disambiguate the particular use.

have on us, does not preclude their having contents that can be propositional in nature. I will point out two primary defects of the causal account proposed by Davidson: one, it combines the theses of literalism and compositionality to the analysis of metaphor and in so doing mischaracterizes metaphors as having only literal meanings;[2] and two, it presents a one-sided perspective on the use of metaphors in terms of the effects they have on hearers, and, thereby, fails to appreciate the value in the making of metaphors when hearers become speakers. The use of metaphors in disagreements, deductive and inductive arguments, and the possibility of retracting metaphorical utterances, making of inferences from metaphors—all these practices establish one crucial thing: contra Davidson, metaphors could have contents that are propositional in nature.

2. AGAINST PROPOSITIONAL CONTENTS OF METAPHORS

The causal theorist (Davidson, 1979; Rorty, 1979; 1987; 1989; Reimer, 2001; Lepore & Stone, 2010) is motivated to restrict the semantic notions of meaning and truth to the more familiar literal uses of language. She is averse to both revising her ontological commitments, and broadening the use of truth and meaning, to include, or apply to, metaphorical and other figurative uses of language. Literal uses of language can be evaluated for truth partly because there are generally accepted ways for fixing the contents and propositions expressed by literal sentences (or utterances), and partly because, unlike in the case of metaphorical sentences, literal truth conditions, usually, can be assigned to sentences irrespective of the particular contexts in which they are used. Every metaphorical claim or sentence, when construed literally, has a literal content or expresses a literal proposition. The causal theorist is of the view that the literal content or the proposition the metaphor literally expresses is the only content possessed or proposition expressed by a metaphor; the non-literal aspect of a metaphor is nothing propositional. This view implies that metaphors do not have propositional contents in addition to their literal contents, and hence, metaphorical sentences *qua* metaphors cannot be truth-evaluable. This presents a bit of a puzzle: on the one hand, in virtue of being a metaphor, a metaphorical sentence is meaningful and has a non-literal content, and yet the metaphor itself is non-truth-evaluable; and

[2] I have pointed out in a previous paper (Kwesi, 2018a) that the theses of literalism and compositionality are often linked with another thesis, representationalism, to support the view that the relevant criterion of truth is the capacity to represent states of affairs as they really are. For instance, Cooper, a defender of Davidson's view, has made these remarks: "The notion of truth, as we normally understand it, is used to appraise utterances in terms of what they achieve. A true statement is one which successfully achieves what statements generally aim to achieve—telling how things really are. To employ the notion of truth in the appraisal of metaphor, therefore, wrongly suggests that metaphors, too, have the dominant aim of getting us to see how things actually are" (1986, p. 250). I argued there that the combined theses of literalism and representationalism do not support the view that metaphors are not truth-evaluable.

on the other hand, a metaphor has or expresses only a literal proposition which makes the metaphor either literally true or false, and thereby, making the metaphor truth-evaluable. The causal theorist argues for the second part of the puzzle by showing that

i. the words of a metaphor have only literal meanings, and, therefore, the metaphorical sentences they compose only have literal meanings; and

ii. in light of (i), metaphorical sentences have literal truth conditions which makes most, if not all, metaphors patently or literally false.

For the first part of the puzzle, she supports her position by arguing that

iii. the supposed additional non-literal meaning or content of a metaphor is not propositional in nature;

iv. this non-propositional meaning of the metaphor is merely the effects metaphors have on their recipients; and

v. a metaphorical sentence does not have a single definite content or meaning; rather, it has many, and perhaps, an infinite number of contents.

Davidson's anti-truth account of metaphor gives expression to the tenets (i)—(v) above. Davidson's (1979) main claim (as he himself calls the "thesis" of his paper) is that "metaphor means what the words, in their most literal interpretation mean, and nothing more" (p. 30). This thesis is borne out of a commitment to two views about language: literalism and compositionality. Davidson's literalism acknowledges a distinction between the literal and metaphorical uses of language, but claims that sentences can only have ordinary literal meaning and truth, and that a distinction between the literal and the metaphorical does not entail that metaphorical sentences have "special" meaning and truth in addition to their literal senses and truth. What metaphors mean, and what their truth values are, are no different from their assessment from a literal point of view. In his commitment to compositionality, Davidson is of the view that the meaning of a sentence is determined from the meanings of the individual words that compose it. If a metaphor can only be explained by appealing to the literal meanings of the words that compose it, then for Davidson "sentences in which metaphors occur are true or false in a normal, literal way, for if the words in them don't have special meanings, sentences don't have special truth" (p. 39). Combining his literalist and compositionalist views, Davidson's claim is that the words of a metaphorical sentence have no special meanings other than their ordinary literal meanings, and hence the sentences they compose only have literal meanings.

In view of the fact that metaphorical sentences only have literal meanings and literal truth conditions, metaphorical sentences have no contents except the contents that they literally express. This is why most metaphors are literally false, if not absurd. That metaphors have no contents (except what they literally express)

implies that there is nothing else that is communicated or conveyed by the use of metaphor, nothing else propositional that can be grasped and evaluated as true or false. Davidson entreats us to give up "the idea that a metaphor carries a message, that it has a content or meaning (except, of course, its literal meaning)" and see the supposed content of metaphor as "something about the *effects* metaphors have on us" (p. 43). A metaphor can provoke thoughts and ideas in us, it can make us attend to some likeness and similarities between two things, it can cause us to notice something in a different way, but all these are effects metaphors have on us: metaphors "make us appreciate some fact—but not by standing for, or expressing, the fact" (p. 44). Davidson's denial of the cognitive claims of metaphor presents us with an account of metaphor that is causal in nature: it is a causal account in the sense that it explains metaphor both in terms of what it causes us to do and the effects it has on us. In this cause-effect view, metaphor has no content other than what it literally means and expresses, which is usually false or absurd; if we mistakenly think that there is an additional figurative or metaphorical content to a metaphor, it is merely because we are confusing effect with content. What metaphor directs our attention to, what it makes us see, cannot be propositional in character; for as Davidson exclaims: "seeing *as* is not seeing *that*" (p. 45). In this regard, Davidson likens metaphor to a joke or a dream or "a bump on the head"—these acts have effects on us by making us come to notice or observe some fact without their expressing those facts. Metaphors can lead one to see something as, but not that; they can intimate, nudge, or poke one to view something in a different way, but intimation is not the same as meaning; they can cause one to have certain beliefs, but they do not express those beliefs;[3] like jokes and bumps on the head, they can have effects on others, but such effects are not propositional elements that can be evaluated on the basis of semantical notions like meaning, truth and reference.

Davidson argues also that our inability to paraphrase or decide exactly what the content of a metaphor is, is not primarily because metaphors are non-paraphraseable, but because there is no content to be paraphrased or expressed. He thinks that we imagine there is a content to be captured when all the while we are in fact focusing on what the metaphor makes us notice; we are merely focusing on the effects metaphor has on us. He writes:

> If what the metaphor makes us notice were finite in scope and propositional in nature, this would not in itself make trouble; we would simply project the content the metaphor brought to mind onto the metaphor. But in fact, there is no limit to what a metaphor calls to our attention, and much of what we are caused to notice

[3] By causing us to form certain beliefs that such-and-such is the case, there is something "propositional" about metaphor, that is, the acquisition of propositional attitudes. But the Davidsonian contention is that the metaphorical sentence *itself* does not express the proposition that such-and-such is the case; the metaphor itself does not make a statement or communicate something that is propositional.

> is not propositional in character. When we try to say what a metaphor "means", we soon realize there is no end to what we want to mention. (1979, p. 44)

Davidson's line of thought has been expanded by Lepore & Stone (2010) in their thesis statement that "though metaphors can issue in distinctive cognitive and discourse effects, they do so *without* issuing in metaphorical meaning and truth, and so, *without* metaphorical communication" (p. 166). Like Davidson, they take a pragmatic view of metaphor as involving some sort of speaker intentions and not c o m m u n i c a t e d m e a n i n g. They argue that metaphor should be catalogued among practices such as "hinting, joking, trash-talking, flirting, and flattering" (p. 166). By joking, one aims to cause certain effects in one's audience rather than to assert something that can be appraised for truth. And metaphor is no different from jokes. An interlocutor may use a metaphor with the intention that his hearers see a particular point but this point "is not a property of the metaphor itself" (p. 173). Lepore and Stone contend that "interlocutors use their metaphorical discourse not to assert and deny propositions, but to develop imagery and to pursue a shared understanding" and that "such practices can account for our interactions in using metaphor, without appealing to metaphorical meaning or metaphorical truth" (p. 177). In effect, they argue for distinguishing "metaphorical thinking—developing imagery, seeing one thing as another, noticing similarities—from merely grasping a proposition, namely the one that is speaker meant, brought about through an intention to present information through coordination or intention recognition" (p. 178).

For Richard Rorty, Davidson's causal account enables us to see the distinction between the literal and the metaphorical not as two sorts of meaning or truth but a distinction between "familiar and unfamiliar uses of noises and marks" (1989, p. 17). The literal is the regular and familiar uses of language that are marked by predictability and a generally accepted procedure for determining meaning and truth. The metaphorical, Rorty thinks, is an unfamiliar noise—a use of familiar words in unfamiliar ways. As an unfamiliar noise, metaphor has no fixed place in the language game. Uttering a metaphorical sentence is not to say something true or false; it is not to say something that has a meaning. Rather, uttering a metaphor only produces an effect in one's audience and causes them to have certain beliefs or act in certain ways. In one characterization of metaphor, Rorty has this to say:

> Tossing a metaphor into a conversation is like suddenly breaking off the conversation long enough to make a face, or pulling a photograph out of your pocket and displaying it, or pointing at a feature of the surroundings, or slapping your interlocutor's face, or kissing him. Tossing a metaphor into a text is like using italics, or illustrations, or odd punctuation or formats. All these are ways of *producing effects* on your interlocutor or your reader, but not ways of conveying a message. To none of these is it appropriate to respond with "What exactly are you trying to say?" If one had wanted to say something—if one had wanted to utter a sentence with a meaning—one would presumably have done so. (1989, p. 18, italics mine)

In another characterization, Rorty likens metaphor to thunderclaps and birdsong to make the same point. A novel metaphor is like the noises of a bird we are not acquainted with. The noise c a u s e s us to believe that there is, for instance, a quetzal in the forest. But the noise itself does not convey the information nor express the fact that there is a quetzal in the forest. In the same way, a metaphor causes us to change our beliefs and desires without representing or expressing any facts of the world. According to Rorty, we should see metaphor in its functions as

> causes of our ability to do lots of other things—e.g., be more sophisticated and interesting people, emancipate ourselves from tradition, transvaluate our values, gain or lose religious faith—without having to interpret these latter abilities as functions of increased cognitive ability. (1987, p. 284–285)

Rorty, therefore, allows metaphors to have functions, that is, to be causes of beliefs, just as Davidson endows metaphor with the ability to direct our attention to notice similarities between things. Yet, these functions of metaphor are not to be interpreted as conveying any message that will add to our knowledge.

Both Rorty and Davidson rely on a distinction between "cause of belief" and "justification of belief" (or "reason for belief") and argue that it is a conflation of this distinction that seems to give some credence to the cognitive claims of metaphor. As it pertains in sensory observations of birdsong and other unfamiliar noises, we can draw a distinction between the unfamiliar noise as a stimulus to knowledge and the claim that it conveyed that knowledge. The noise is merely a stimulus to knowledge or a cause of the belief that there is a bird in the forest, but it is not a reason for, nor a justification of, the belief or information that there is a bird in the forest. What causes belief and knowledge is not necessarily that which expresses or conveys belief and knowledge. Metaphor as an unfamiliar noise belongs not to cognition but to stimulus. It has a place in a causal scheme of things, but it does not have in addition a place in a pattern of justification of beliefs. By confining the interpretation and meaning of metaphor to the literal and explaining away the supposed additional content of a metaphor in terms of the effects metaphors have on us, Davidson, Rorty, Lepore and Stone, and others, limit the semantic notions of truth, meaning and content to regular and literal uses of language.

3. Analysis and Critique of the Causal Account of Metaphorical Content

Davidson's account of metaphor has been discussed extensively in the literature, mostly in two main directions: there are those who criticize his literalist account and argue for the cognitive claims of metaphor, especially Black (1979; 1993), Goodman (1979), Leddy (1983), Hesse (1987; 1988), Farrell (1987), Moran (1989), Camp (2006a; 2006c; 2008), Johnson (2008), and most works in cognitive linguistics; and there are others who have defended his account, partic-

ularly, Davies (1982), Davies (1984), Rorty (1987; 1989), Cooper (1986), Reimer (2001), Lepore & Stone (2010; 2015). The critique of the Davidsonian account in the literature has primarily focused on showing that there is a cognitive dimension (Black, 1979) or a propositional dimension (Moran, 1989; Camp, 2006a; 2006b; 2008) to metaphor. This propositional aspect of a metaphor is usually derived from, or associated with, the intentions of the speaker—what the speaker means by uttering a metaphor (Searle, 1979).

The critique of the account I offer here is partly diagnostic, intended to reveal the ways in which the Davidsonian tenets (i)—(v) above are flawed and untenable; and it is partly prescriptive, meant to provide evidence that metaphorical sentences have propositional contents. The analysis pursued here is to show how Davidson's account does not adequately and satisfactorily deal with the issue of the meaning and content of metaphor. It is often regarded as implied by Davidson's account that once one accepts his central thesis that a metaphor has no additional meaning and truth-value other than its literal meaning and truth-value then one is committed to seeing metaphor only in terms of its functions—in terms of its causes and effects. However, the inference from literalism—"only literal meaning"—to a causal explanation—"only causal role"—is not a logically necessary one. It is possible to accept Davidson's central thesis without adducing a causal explanation for how metaphor works, and more importantly, without singling out a causal explanation as the only explanation one could give to metaphor. Similarly, the conclusion that metaphors have no propositional contents cannot be premised on the fact that metaphors have causes and effects on their users. It is possible to accept a causal explanation of how metaphors work—that is, that they cause us to acquire certain beliefs, that they direct our attention to see similarities between two things, etc.—and posit that they have propositional contents in addition to their causal role. In other words, that metaphors have causes and effects does not preclude their having propositional contents. Metaphors do have functions, they do cause us to do certain things, they have effects on us; but their having functions and effects is not a reason for, nor a limitation of, their capacity to be something else, or have something more—something propositional.

3.1. Metaphor and Compositionality

Davidson's ultimate position on metaphor is that metaphorical sentences have only literal meanings and hence only literal truth conditions. This position is as a result of combining his thesis of literalism—that the words of a metaphor have only literal meanings—with compositionality—that the meaning of a sentence is derived from the meanings of the individual words that compose it. That is, if the words of a metaphor have only literal meanings then metaphorical sentences have only literal meanings. However, this analysis is flawed: the mistake lies in the conjunction of the two theses—literalism plus compositionality—to generate the solution that metaphorical sentences have only literal meanings:

Literalism (L) + Compositionality (C) = Literal Meaning (LM)

To see the flaw, we have to take a critical look at the two principles of literalism and compositionality. Compositionality is regarded as one of the essential properties of language which is used to explain, among other things, our linguistic and cognitive abilities to learn a language by learning the meaning of a finite number of expressions and yet be able to produce and understand an infinite number of meaningful sentences (Davidson, 1967, 1984; Fodor & Lepore, 2002; Pagin & Westerstahl, 2010a; 2010b). On one definition, the principle of compositionality is the claim that "the meaning of a complex expression is determined by its structure and the meanings of its constituents" (Szabo, 2010, p. 255). This determination of the meaning of the complex expression is usually construed in functional terms; that is, "the meaning of the complex expression is a function of the meanings of its parts and the mode of composition by which it has been obtained from these parts" (Kracht, 2011, p. 57). Compositionality is a semantic phenomenon, for it determines the semantic value of a complex from the values of its constituents, thereby constraining the relevant factors involved in the determination of meaning. In a strict sense of compositionality, what is necessary and sufficient for determining the semantic value (meaning, content, denotation) of a complex expression is the semantic information and contribution derived from the parts of the complex expression and its mode of composition. This is akin to what Dever (2008) has called the "semantic closure" constraint of the principle of compositionality.

However, it has been questioned in various ways in the literature as to whether the meaning or content of a complex expression is determined purely from the semantic values of, or the semantic information provided by, its constituents and their mode of composition. This questioning arises out of the observation that the meaning or content of an expression is underdetermined by the semantic information provided by the parts of the expression, and that, there are certain constituents of the meaning of an expression that are provided purely on pragmatic grounds, usually by a process of "free pragmatic enrichment" (Carston, 1988; 2002; Recanati, 2004; Sperber & Wilson, 1995; Hall, 2009). The utterances of "I have had breakfast" and "It is raining" have their truth-evaluable contents <I have had breakfast *this morning*> and <It is raining *in Cape Town*> respectively, where the time (of breakfast) and the location (of rain) are freely pragmatically supplied by the context of the utterance. These additional constituents of the meaning or content of an expression are not traceable overtly or covertly to the encoded meanings of the parts of the expression; they are provided and constrained by purely pragmatic factors.[4] Generalizing from this observation, con-

[4] Some semanticists like Stanley (2000; 2002), and King & Stanley (2005) have argued that indexicality and other contextual factors can be traced to the logical form of the expressions which suggests that the so-called free pragmatic enrichments are constrained semantically. Lasersohn (2012) has argued that the context-sensitivity nature of most expressions and the fact that speakers rely on pragmatics to arrive at the contents of cer-

textualists and pragmatists argue that the intuitive meaning and content of an expression cannot be given solely by a compositional semantics.

Compositionality itself as a principle for the determination of the semantic value of a complex expression does not discriminate between literal and non-literal meaning even though it seems to presuppose literal meaning. All that is required for compositionality is that the meaning of the complex be a function of the meanings of its parts and their mode of composition. This does not imply that the kind of meaning[5] of the complex be determined from a function of its parts and the ways in which they are composed. That is, compositionality does not specify or stipulate that the meaning-type of the complex be derived from those of its units. The requirement that the meanings of complex expressions be literal because their units are literal is an additional constraint on the meaning of complex expressions. Up to this point we had noted that compositionality requires that the meaning of the complex expression is a function of both

(a) the meanings of its constituents, and

(b) their mode of composition.

Now, there is a further constraint on the meaning of complex expressions or sentences in general in relation to the parts that compose them:

(c) the kind of meaning (or the meaning-type) of the complex expression is a function of the meaning-types of its constituents.

This additional constraint (c) is what informs the literalism of Davidson's account of metaphor.

We should distinguish between two strands of literalism: word-literalism (L_w) and sentence-literalism (L_s). Davidson actually argued for word-literalism, indicating that words themselves do not have "extra" or non-literal meanings, and by extension, he argued for sentence-literalism through a compositional analysis. The fact that Davidson argued for word-literalism has been observed also by Farrell (1987) who shows that in his essay on metaphor, Davidson hardly treats metaphor at the sentential level; rather, he treats it at the level of word meaning. At the beginning of his paper where he states his main thesis, Davidson writes: "This paper is concerned with what metaphors mean, and its thesis is that metaphors mean what the *words*, in their most literal interpretation, mean, and nothing more" (1979, pp. 29–30, emphasis mine). He again writes that "my disagreement is with the explanation of how metaphor works its wonders. To anticipate: I depend on the distinction between what words mean and what they are

---

tain expressions do not undermine, but are compatible with, the principle of compositionality.

[5] By "*kinds* of meaning" I mean meaning-types such as literal meaning, metaphorical meaning, symbolic meaning, etc.

used to do" (p. 31). And when he discusses and rejects other views on metaphor, Davidson has these things to say: "The idea, then, is that in metaphor certain words take on new, or what are often called "extended" meanings" (p. 32); "Perhaps, then, we can explain metaphor as a kind of ambiguity: in the context of a metaphor, certain words have either a new or an original meaning" (p. 32); "I have been making the point by contrasting learning a new use for an old word with using a word already understood" (p. 37). And after discussing a number of views, Davidson concludes:

> The argument so far has led to the conclusion that as much of metaphor as can be explained in terms of meaning may, and indeed must, be explained by appeal to the literal meanings of words. A consequence is that the sentences in which metaphors occur are true or false in a normal, literal way, for if the words in them don't have special meanings, sentences don't have special truth. (1979, p. 39)

As Farrell points out, Davidson's focus on the word instead of the sentence "serves his strategy in the article" for "he interprets his opponents to be making a claim that metaphorical meanings constitute an extra layer of *word* meanings, and consequently, that metaphor is analogous to ambiguity, in that a word may have two different meanings" (1987, p. 637).[6] So, Davidson inveighs against positing additional metaphorical or figurative meanings to the words that compose a metaphor and then argues that since the words do not have extra meanings other than what they mean literally, metaphorical sentences only have literal meanings. In doing so, Davidson rescues semantics from accounts based on extended word-meanings and also from the multiplicity of meaning and truth with respect to the words in a metaphor. Now, so long as we are dealing with the words of a metaphorical sentence, Davidson's account seems plausible.

A point of departure with Davidson here is that both his attack on the theories, and the theories he was attacking, miss one crucial point about metaphor: a metaphor is not necessarily a metaphor because a word has been used metaphorically or in an unfamiliar way. It is only when we take the word, be it the focal word of the metaphorical sentence, as the unit of analysis that we worry as to whether the word has an "extended" meaning or reference. Indeed, words in every sentence have no "extra" meanings other than what they mean literally, but their composition into sentences marks an important difference between figurative and literal sentences. That is, word-literalism does not imply sentence-literalism when the expressions in question have been construed metaphorically or figuratively. It is one thing to say that the words in a metaphor only have literal meanings and

[6] In Davidson's general theory of meaning, however, word and sentence meaning go hand in hand. For him, other than a semantic theory being compositional, it must also be interpretative, in the sense that it should be possible for the theory to be used to understand speakers and their linguistic behaviour. Since a theory of meaning is a theory of truth for Davidson, one constructs a systematic truth theory from both the meanings of the words and sentences of a language.

another thing to say that the metaphorical sentence has only a literal meaning or interpretation. One can endorse the claim that the words of a metaphor have no special, extra, non-literal meanings without further endorsing the claim that the metaphorical sentences composed out of the individual words have literal meanings. Idiomatic expressions are paradigmatic cases of counter-examples not only to compositionality in general but more particularly to the constraint (c) on the meanings of complex expressions or sentences in general which requires that the meaning-type of the constituents transfer to the meaning-type of the complex expression.

Idioms are generally considered to be expressions whose meanings are conventionalized in the sense that "their meaning or use can't be predicted, or at least entirely predicted, on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another" (Nunberg, Sag, & Wasow, 1994, p. 492). An idiomatic expression defies the principle of compositionality in that the meaning of the idiomatic expression is not determined by a compositional function of the meanings of its constituents (Chomsky, 1965; 1980; Katz, 1973; Kracht, 2011).[7] The meanings of idiomatic expressions like "kick the bucket" and "take the bull by the horns" are not determined by the meanings of their component parts despite their having syntactic structures. Interestingly also, the words of these idiomatic expressions do not acquire extra meanings other than their literal meanings, but the idiomatic meanings of the expressions are not dependent on the literal meanings of the words even where their composition into a whole fails. That is, in spite of the fact that the parts do not compose into a whole in determining their idiomatic meanings, the idiomaticity of the expressions is not a function of the idiomaticity of the words that make them up.[8] In other words, if the constraint of the meaning of complex expressions (c) holds, then when the expressions are given idiomatic meanings this should result from the constituent words having idiomatic meanings (just as when they are interpreted literally, the words should have their literal meanings at play). But although the expressions have idiomatic meanings their constituent words do not acquire any extra meanings other than their literal

[7] Nunberg, Sag, & Wasow (1994) have shown that not all idiomatic expressions are non-compositional. They distinguish idiomatically combining expressions like "pull strings" whose meanings could be distributed among its parts, from idiomatic phrases like "kick the bucket" which do not distribute their meanings to their parts. My concern in the main is with idiomatic phrases.

[8] One could point out that this is so because idioms are lexicalized expressions that should be treated as single words. However, this view is unattractive. As Titone and Connine point out, "there is evidence showing that idioms possess a great deal of internal sematic structure. Idioms are modifiable with adjectives or relative clauses (e.g. "She did not spill any of those precious beans"), and parts of idioms may be quantified (e.g., "She didn't spill a single bean"), emphasized through topicalization (e.g., "She didn't spill the beans yesterday, but spilled them today") without disrupting comprehension or awareness of their idiomaticity" (Titone & Connine, 1999, p. 1659).

meanings, and hence, the constraint on the meaning of complex expressions (c) cannot be accurate.

The point here is that we can allow that the words that compose any figurative expression maintain their literal meanings, but this concession does not imply that figurative expressions only have literal meanings. This is because the analysis and interpretation of a figure of speech like an idiom starts rather at the phrasal or sentential level. An idiom obviously is different from a metaphor—a metaphorical expression can be live and novel in characterizing one thing in terms of another thing, while an idiom is a set phrase whose meaning cannot be inferred from the meanings of the words that make it up, and whose usage is characteristic of a group of people. A significant difference between a metaphor and an idiom is that unlike an idiom, an understanding of the literal meanings of the words in a metaphor aids in the interpretation of the metaphor. However, the analysis and interpretation of metaphor takes a cue from idioms: a metaphorical sentence can have a meaning, a meaning other than what it literally means even though the words that compose the metaphor as Davidson has strongly argued only have literal meanings.[9] How does this cash out?

In any context of use, both metaphorical sentences and idiomatic expressions can be understood and interpreted literally. Compare the idiom "she kicked the bucket" to the metaphor "Gabriele is a crocodile." When we combine the literalist thesis with compositionality "she kicked the bucket" just means that she kicked the bucket, and similarly, "Gabriele is a crocodile" means that Gabriele is a crocodile. Construed figuratively, it seems okay to say that "she kicked the bucket" means that she is dead. Or perhaps, we should say that in an appropriate context, one utters "she kicked the bucket" to mean that she is dead. "That she is dead" becomes the content or the proposition asserted by the idiom-user. (This is quite different from the effect the idiom might have on an audience, if any.) If the sentence "she kicked the bucket" could mean both she kicked the bucket and she is dead then we can say that the sentence has two meanings depending on the use to which it is put: used literally, it has the meaning (LM) that she kicked the bucket, and used figuratively (as an idiomatic expression), it has the meaning (MM) that she is dead. The difference between LM and MM lies in the role

[9] It is possible for one to argue that idiomatic expressions are not necessarily breaches of compositionality, and that the cases of idioms neither affect nor make compositionality false. The rule of compositionality is meant to apply to non-idiomatic uses of language. This argument seems right. But the point here is not that the rule of compositionality is breached or made false by idiomatic expressions. The point rather is that idiomatic expressions, in being figurative expressions, do not require that their meanings imply that the words that make up the expressions also acquire figurative meanings. This suggests that the meaning of a figure of a speech does not imply that the words of that figure of speech have figurative meanings or applications. So, in the case of idioms, their meanings do not depend on the meanings of the words that compose them, and there is no further requirement that the words should have figurative meanings. This is the principle that I am claiming holds in the case of metaphors.

compositionality plays in the determination of the meaning of the sentence: whereas LM results from the meanings of the individual words of the sentence, MM does not; MM is not worked out from the meanings of the individual words of the figurative expression.

A similar situation is what obtains in metaphor. In terms of LM, "Gabriele is a crocodile" means that Gabriele is a crocodile, which might seem false or absurd. But understood figuratively, it can have the MM meaning that Gabriele is impulsive and angry. What is interesting about the metaphorical case is that the MM meaning, while it does not result out of the composite of the literal meanings of the words (for then we will have LM), is linked in a peculiar way to the words of the sentence, not in terms of literal meaning, but usually, in terms of certain cultural and idiosyncratic features or connotations associated with the words of the metaphor. These cultural, religious, moral, aesthetic and idiosyncratic features we associate with certain words and phrases are, in an important sense, not part of what we will ordinarily call the literal meaning. It is not part of the literal meaning of Gabriele being a crocodile that he is impulsive and angry. This suggests that a determination of the meaning of the metaphor from a composite of the literal meanings of the words will mischaracterize the expression as a metaphor. In both the idiomatic and metaphorical cases, the words of the expressions retain their usual literal meanings, but a compositional determination of their meanings misses the point of their figurativeness; that is, apprehending their LM meanings is just to take the expressions literally.

The idiom, "she kicked the bucket" means she is dead, period. However, if we understand both metaphors and idioms as figurative devices which defy the laws of compositionality and constraint (c) resulting in metaphorical and idiomatic expressions having MM as characterized above, then it seems that we cannot think of the MM of a metaphor as merely a paraphrase (or effect or insight) and that of an idiomatic expression as a meaning or proposition. It is true that there could be more than one interpretation or meaning we could come up with for a metaphorical sentence; but this will not yield different kinds of meanings. There are only the two kinds of meanings here—LM and MM—depending on whether the meanings are calculated based on compositionality or not. Just as a literal sentence could be ambiguous or have multiple meanings under LM, so a metaphorical sentence could have a variety of meanings under MM; the various meanings under MM are all possible meanings that are partly determined and constrained by the contexts and circumstances in which the metaphorical sentence is used. The distinction between LM and MM in terms of whether they are faithful to the principle of compositionality can be used to mark a difference between metaphors and ambiguous sentences. The different meanings of an ambiguous sentence are all determined by a compositional analysis of the literal meanings of the words where either the different lexical meanings of the words are used in the analysis (as for instance in the case of "he went to the bank") or that the compositional structure is permuted (as in the case of "he killed the man with an umbrella"). Metaphors are not ambiguous either lexically or structurally.

For the different meanings of a metaphor belong to MM which does not entertain the use of compositional analysis.

So, in agreement with Davidson, the words in a metaphor, like those of most figurative expressions, retain their ordinary meanings and significations, but unlike Davidson, the meaning of a metaphorical sentence is not computed from the literal meanings of the words that make it up. Metaphors and other figurative expressions defy the principle of compositionality. Since the principle of compositionality does not apply in the case of metaphors, it implies that metaphorical sentences, contra Davidson, do not have only literal meanings and should not be evaluated with literal truth conditions. Does this imply that metaphorical sentences have additional meanings other than their literal meanings? This will amount to similarly asking whether an idiomatic expression has an additional meaning other than its literal meaning. Is the meaning that she is dead an additional meaning of the idiomatic expression "to kick the bucket"? It is obvious that the idiomatic meaning of "she kicked the bucket" just *is* that she is dead, because the sentence has been construed figuratively or idiomatically. And as we have seen, this meaning is the MM that is not a resultant of the compositional analysis of the words of the sentence. This MM is not a meaning in addition to the LM of the sentence, since the sentence has been construed figuratively. In the same vein, construing the sentence "Juliet is the sun" literally and realizing that the sentence is false or absurd is not an indictment on the sentence when it is construed metaphorically. The MM of the metaphor is not a meaning extra or additional to its LM, as if they are derived from the same analysis. What exists here is a meaning difference in kind, which reflects a difference in construal of the sentence: a sentence construed literally employs a compositional analysis in determining what it means literally; that same sentence construed metaphorically or figuratively, adopts a non-compositional analysis in determining what it means non-literally.

In summary, we have three models for associating meaning and content with metaphors in relation to the literal:

Table 1

Model	Word	Progress	Sentences
Model 1	Literal meanings Metaphorical meanings	Compositionality →	Literal meanings Metaphorical meanings
Model 2	Literal meanings	Compositionality →	Literal meanings
Model 3	Literal meanings Literal meanings	Compositionality → Non-compositionality	Literal meanings Metaphorical meanings

The first model is the view that Davidson attacks which posits that metaphors trade on the ambiguity of words, that words have literal and metaphorical senses, and hence metaphorical sentences have two meanings (Beardsley, 1962; 1978; Goodman, 1968; 1979). This model ought to explain how words come to acquire metaphorical meanings and how they are composed to form metaphorical meanings of sentences. A major problem for this model is how it can satisfactorily explain the phenomenon of dead metaphors: as Davidson puts it, when the dead metaphor "he was burned up" was active, "we would have pictured fire in the eyes or smoke coming out of the ears" (1978, p. 38). Davidson's own view is the second model which posits that words have only literal meanings and the sentences they compose also have only literal meanings. But this view is not able to satisfactorily explain how one comes to fail to grasp the metaphor even though one understands its literal meaning; and as I will show below, the view is also not able to account for how two people can disagree over the proposition expressed by a metaphor even in situations where the literal meaning of the metaphor seems irrelevant to the disagreement or where the two parties can engage in disagreements even though they agree on the literal meaning of the metaphor.

The view I have tried to formulate above is the third model which grants that words in a sentence have only literal meanings, but the meanings of the sentences they constitute are either literal or metaphorical depending on whether the meanings are derived from a compositional analysis or not. The affinity of metaphors with idioms that I drew above suggests that the non-compositional transition from literal-word-meaning to metaphorical meaning is a matter of sentence, as opposed to speaker, meaning. Speaker-meaning is a viable alternative route that is non-compositional in nature, but this route need not make any pronouncement about the literal-metaphorical status of the words of a metaphor. The third model pursued here is an affirmation of the fact that the words of a metaphor do not acquire mythical or mysterious non-literal meanings. But more im-

portantly, one can use an idiom to speaker-mean (SM) something other than its LM or MM; SM is pragmatic meaning and it arises in the use of both literal and figurative expressions and in our linguistic practices in general. This means that for literal sentences, we can have LM and SM as in the case of implicatures; and for figurative expressions we can also have LM and SM, as for instance in the cases of understatements and ironies. But SM alone does not establish that figurative expressions like idioms have, in addition to LM, MM which is semantic meaning. The model pursued here is that in addition to SM and LM, metaphorical expressions have MM, and this semantic meaning can be different from, or similar to, although not necessarily derived from, the pragmatic speaker-meaning.

3.2. The problem of Many Contents

A possible objection to the analysis above is that there is a kind of *definiteness* associated with literal meaning and content such that even if we allow both metaphors and idioms to have MM, that of the idiom is definite and given. We cannot appropriately talk of *the* meaning or *the* content of a metaphor as we do with an idiomatic expression. If there is no definite content to a metaphor, this will suggest that it is not a genuine linguistic item that we should be concerned with associating it with meaning and content. However, this objection is not well motivated. The point of the "inexhaustibility" (Cohen, 1975) of the interpretation of metaphor cannot, and should not, be construed as a defect of metaphor. It should also not be construed as the yardstick for attributing content to metaphor. Inability to paraphrase a particular metaphor and/or the indeterminacy of the right kind of paraphrase for a metaphor, are not in themselves indicators of the absence of any content that the metaphor might have. Rather, the ability to paraphrase (most) metaphors into propositional form is an indication that metaphors have contents.

If a metaphor expresses two or more propositions or if it has more than one interpretation, or if it can be paraphrased into more than one sentence, then it is not a matter of its having no content but that it has "many contents". A denial of the content of metaphor rests on the flawed principle that many contents mean no content at all; it is like when you say too much, you end up not saying anything at all. Although a metaphor says too much, it at least says something. And it is because it says something that we are able to give at least one paraphrase of it.

The objection that many contents imply no genuine linguistic item loses its sway when we consider treatments of vagueness and borderline predicates in the literature where vague sentences are made truth-evaluable by such methods as supervaluationism (Fine, 1975; Keefe, 2000; 2008; Cobreros, 2008). Vagueness is considered a semantic phenomenon[10] (Keefe, 2008; Cobreros, 2008) resulting

[10] Williamson (1994) for instance, regards vagueness as an epistemic phenomenon by treating the proposition a vague sentence expresses as a borderline case which is either true or false, but we are ignorant of which value it is.

from semantic indecision in the sense that "nothing in the world, either in the use or in any other factor relevant to the determination of the meaning of a vague predicate, *decides* which of the ways in which we could make precise the predicate is correct" (Cobreros, 2008, p. 292). Vague sentences are therefore considered to be indeterminate; they are neither true nor false. However, a supervaluational model can be applied to a vague sentence to make it either (determinately) true or false by means of an admissible precisification whereby the sentence is made more precise. In this way, the vague sentence is true if and only if it is true on all ways of making it precise, and false if and only if it is false on all ways of making it precise, and neither true nor false otherwise (Fine, 1975; Keefe, 2000; 2008). The point of supervaluationism in relation to vagueness is to show how a vague sentence or a multiple-referring expression can be made truth-evaluable and be accorded a definite truth value. If we can provide a semantics for vague sentences, then, despite the differences between metaphors and vague sentences which there might be, the indefiniteness objection to metaphors cannot be used to deny its capacity to be appraised for truth. For, on a supervaluational operation on a metaphor, one can take a metaphor to be true or false on all admissible ways of precisifying it, where the precisification could be in the form of literalizing or paraphrasing the metaphor. The claim here is that metaphors, like vague sentences and borderline cases, have contents which admit of many possible precisifications/paraphrases; and just as the many contents a vague sentence may have do not preclude it from being appraised for truth, it cannot be correct that metaphors' having many contents implies that they are not genuine linguistic items that can be truth-evaluable.

3.3. Metaphor from the Perspective of the Metaphor-Maker

Besides the rather disparaging remark about metaphor as a noise, both Rorty and Davidson explain metaphor with respect to the effects it has on the hearer. While this may be true—that is, metaphors have certain effects on hearers—the explanation is one-sided and inadequate: on the one hand, it gives no explanation of metaphor from the speaker's perspective; and on the other hand, the explanation it gives cannot effectively be extended to the maker of a metaphor. How is the speaker to understand her metaphorical utterance if she is merely making an unfamiliar noise? What effect is metaphor to have on the maker of the metaphor? How is the metaphor to cause a change in beliefs and desires if it is to be construed as merely a noise from the speaker's perspective? Talk of the effects of metaphor seems accurate when we are considering the role of metaphor from the point of view of the audience or hearer, but it seems inappropriate to suggest that the metaphor also causes certain effects in the one making it. The causal account of metaphor fails to note that there can be both internal and external noises. External noises may have effects on us and cause us to do certain things or behave in certain ways; but internal noises are internally generated, and hence the effects the noises may have on others may not necessarily apply to the generator of the

noise. In the case of metaphor, the effects metaphor is seen to produce do not apply to the maker of the metaphor. Hence, an account of metaphor that only explains metaphor in terms of noises and effects on the part of the audience is an inadequate one.

It is one thing to say that a metaphor can cause one to entertain certain beliefs and propositions, and another thing to say that a metaphor is an outward expression of the beliefs and propositions one has (or is) entertained (or entertaining). We can agree with Davidson and Rorty (for the sake of argument), that the sorts of things that a metaphor may cause one to entertain are not propositional in nature, but this agreement does not imply that the metaphor itself cannot express a proposition that has been entertained by the maker of the metaphor. One cannot use the non-propositional character of the sort of things a hearer is caused to entertain to deny the propositional character of the metaphorical statement that the speaker of the metaphor may assert. What a metaphor may be "used to do", what a metaphor may "cause" one to do or entertain, and the "effect" of what a metaphor may have on anyone, do not offer an analysis of, and cannot be used to explain, what a metaphor *is*.[11] What a metaphor is—a statement or utterance borne out of the beliefs and propositions conceived and entertained by a speaker—and what a metaphor may "suggest" or "point out", are also separate issues. One has to be cautious not to conflate, first, the essence and work of metaphor, and second, the analysis of metaphor from the perspectives of the hearer and speaker. We can delineate the activities of the speaker and hearer of a metaphor from the "work" of the metaphor itself. "Nudging", "poking" and "directing of attention", a metaphor can do, but this work of the metaphor does not say anything about whether metaphors can be associated with the expression of propositional contents. If we are interested in what a metaphor can be used to do, and the causes and effects associated with a metaphor, the analysis can begin from the metaphor itself and the force it has on hearers. If we are interested in the meaning or interpretation of a metaphor, if we are poised to give paraphrases of a metaphor, we can attempt this from the perspective of the hearer by developing strategies and mechanisms the hearer could use, although in most cases, this cannot be done adequately independently of the intentions of the speaker. But we cannot use our conclusions about the causes and effects of a metaphor on the hearer to posit certain assumptions about the making of the metaphor or about the essence of the metaphor itself.

Let us suppose with Davidson and Rorty that we should understand metaphor as *seeing* one thing *as* another thing. From a causal account then, we can explain how a metaphor (or perhaps the metaphor-maker) causes the hearer to see one thing as another thing. But we cannot appropriately explain by the account that

[11] Here, the distinction is between the functions and effects of metaphor on the one hand, and the constitution of metaphor on the other hand. The functions and effects may be used to elucidate what metaphor is, but they cannot stand for what constitutes a metaphor. Similarly, the things a metaphor may cause a hearer to entertain can be distinct from the thing—a proposition, perhaps—a metaphor may assert.

the speaker or the metaphor-maker is caused to see one thing as another thing. In fact, the seeing-*as* experience happens prior to the causes and effects that take place. The metaphor-maker is not caused to see anything by the metaphor and neither does the metaphor have any effect on the metaphor-maker; the metaphor rather reflects what the metaphor-maker has already seen or experienced.[12] A metaphor is like "a bump on the head" Davidson says, but on whose head? It cannot be on the head of the metaphor-maker. A metaphor is like "a joke" but who is to get the joke? It is not about whether the joke is funny or not—a comedian gets his own joke as he is the one making it. A causal account cannot explain the making of novel metaphors even if it can explain the reception of novel metaphors. Thinking of metaphors in terms of effects leaves out the production of metaphors even if the metaphor-maker is using the metaphor to bring out certain effects in others. A cause-effect approach to the understanding of metaphor cannot extend to the making, conception, and evaluation of metaphor.[13]

What causes and effects could there be when the metaphor-maker uses a metaphor in a soliloquy? We can make sense of how a metaphor-maker may attempt to bring about certain effects in his audience, and perhaps, where there is no audience, the intended effect may not be successful or applicable. But this presents a problem in the case of soliloquies where the metaphor-maker is his own audience, that is, in this case, the metaphor-maker utters the metaphor to himself rather than to a perceived audience. In this case, it does not seem right to suppose that the metaphor-maker utters a metaphor to bring about some effects in him or to cause himself to see certain insights. The making of a metaphor is an intentional action and it is not clear how a metaphor-maker nudges himself into noticing things when he utters a metaphor to himself. The causal theorist could ex-

[12] An interesting way of making this point is to take seriously the *is* (identity) of the metaphorical "x is y". The seeing-as view conceives of metaphor as a figure of speech in which the *is* metamorphosis into an *as* so that when the metaphor says that "x is y", it can be understood as seeing x as y. But I as defend in Kwesi (2018b), in metaphor, the metamorphosis is reversed: seeing x *as* y, that is, perceiving or creating a resemblance between x and y, involves a transformative process that changes an *as* into an *is* so that the metaphorical assertion of "x *is* y" is a resultant of the transformative process. The "x is y" is an identity statement, a fusion of the x and y into a new reality. The metaphor-maker asserts that "x is y" not consequently to be caused to see x *as* y, but rather the assertion represents an antecedent seeing of x *as* y that has now transfigured into the claim that x *is* y.

[13] Davidson could respond to the criticism in this paragraph by saying that although the speaker is not caused to do anything, the speaker uses a metaphor with the primary intention of producing certain effects in his hearers. And hence, the making, conception, and evaluation of metaphor can be understood in terms of the intention to produce effects and the subsequent success or effectiveness of those effects on hearers. However, this response is not satisfactory: it merely shifts the locus of the criticism to the effects on hearers. And, as I go on to argue in the next paragraph, this primary intention to produce effects is not applicable in cases where the speaker is speaking to himself alone. In soliloquies, it is not only that the speaker is not caused to make a metaphor, he also does not use a metaphor with the intention of producing an effect in himself.

plain that the soliloquist uses a metaphor *as* if there was an audience and that the absence of an existing audience does not imply that there are no causes and effects of the metaphor; it only shows that these effects do not act on anyone, but would if there were indeed existing audiences. This explanation may seem plausible in a particular kind of soliloquy. We can distinguish between two senses of soliloquies: in the one sense, a speaker makes a speech to himself with an audience in mind like in the cases of practicing a speech one is to give at a later date or when an actor on stage gives a monologue; in the other sense, the speaker has no intended or perceived audiences other than himself as in the cases of thinking out loud or making a note to oneself. It is the second of these senses that causes a problem for the causal account. Where there is a perceived audience, the maker of the metaphor may have certain beliefs about how his metaphorical utterance will affect his audience or the various effects his utterance might have. But the causal account is not able to explain where these beliefs come from and how they are generated. Where there are invisible or perceived audiences, and where there are no audiences at all (perhaps, other than the speaker himself) from the speaker's perspective, the causal account cannot satisfactorily explain how speakers deliberately utter metaphors to themselves.

4. THE SOCIAL PRACTICES OF USING METAPHORS

There are certain features of our use of metaphor that give us good reason to assume that metaphor has meaning and content rather than mere effects on us. Our shared communal practice of employing similar metaphors in everyday discourse attests to the fact that there is a meaning that is grasped and shared by all. Rarely are live metaphors confined to individual speakers in a community. The same active metaphors may be used by a number of speakers or writers in a particular linguistic community. An effect-based approach to metaphor only assumes that one is dazzled upon hearing a metaphor, that one is directed to notice certain similarities between two things. But even if we grant that this is the only business of metaphor—directing one's attention to notice similarities—the ability of two or more people using the same metaphor to enable others to notice the same similarities presupposes that there is a meaning and content of the metaphor that is shared by them.

If Davidson is right that what many people refer to as the content of a metaphor is merely an effect metaphor has on hearers, how can we predict that the same or a similar effect can occur each time a particular metaphor is used? How is the hearer able to grasp a metaphor, exploit it, and use it to produce similar effects on others? How can we judge which effect is appropriate or inappropriate to have in each context of use of the metaphor? If someone is banged on the head but feels no pain, he has a deviant reaction, yet we don't criticize him. But if a hearer fails to get the point of a metaphor—treating it as only having literal content or getting the wrong metaphorical interpretation—then he is apt for criticism. Causal patterns only have deviant instances; and causal deviance doesn't

warrant censure. Since receipt of a metaphor can, on occasion, warrant censure, it is not merely a causal phenomenon. The censure or criticism that is associated with metaphors is even more salient with respect to the making of metaphors. A principal feature of metaphors that was highlighted primarily by rhetoricians is the aesthetic or ornamental value of metaphors: metaphors are useful for embellishing speeches. Hence, rhetoricians developed rules and guidelines for making apt and poetic metaphors that will make speech pleasant. Hackneyed and trite metaphors, and metaphors that involved obscene language were criticized for being unpleasant to the ear, and the makers of such metaphors were seen to lack the artistic skills of making figures of speech. If metaphors are like "bumps on the head" as Davidson argued, the criticisms associated with metaphors will not be applicable; indeed, talk of using metaphors to embellish speech or appreciating metaphors for their aesthetic value will be meaningless. To the extent that some metaphors can be appraised as live, vivid, insightful, astute, and to the extent that some metaphors can be criticized as being banal, pale, unimaginative, metaphors are not merely causal prods; for nudges and pokes and prods are not inherently praised or criticized.

If we can meaningfully talk about grasping or understanding a metaphor, what is it that we grasp and understand? The effect? The content? Isn't the ability to grasp a particular metaphor and effectively use that metaphor in other contexts with the expectation that others understand and utilize that metaphor an indication of something more than effects at play? If metaphor only has a point or if it merely intimates one to see something in a certain way, we cannot conclude from this that grasping the point of a metaphor or being nudged to perceive certain similarities will result in one using the same metaphor to put across the same point or to nudge others to perceive the same similarities. It is very mysterious how one becomes aware of the effect of an utterance on him and whether that was the intended effect of the utterance, and that in using the same utterance he will be bringing about the same effect. Also, if all there is to metaphor is the effects it has on one, it is not clear whether the effects include the ability to pass on the same metaphor to achieve similar effects in others. And similarly, an effect-based account of metaphor cannot explain one's ability to teach and explain metaphors to others, for in teaching and explicating metaphors to others we do not just indicate what the causes and effects of metaphors are. Sameness of the effects of the metaphor on two people cannot account for their ability to use the metaphor to produce the same or similar effects. Meanings are such things that are transferrable; effects are generally not.

The difference in the abilities of two people to use the same metaphor can be attributed to their understanding—grasp of meaning—of the metaphor. To be able to use a metaphor in multiple contexts, to be able to use a metaphor to intend to achieve a different effect, marks the presence of understanding of the metaphor, such an understanding involves both the grasp of meaning and the ability to use the metaphor (Kwesi, 2019b). Meaning is, therefore, at stake in

both the ability to use a metaphor and the inability to use a metaphor in situations where one is aware of the effect of the metaphor.

In practice, hearers rarely ask for the meaning or interpretation of metaphors they freshly encounter; yet, they work out the meaning of those metaphors and employ the same metaphors in their own discourses with the expectation that other hearers will be able to work out what the metaphors mean. The capacity to work out what a novel metaphor means, unaided by the metaphor-maker, involves, at the very least, a kind of reflective comprehension of the effects of the metaphor. Hearers can become users of certain metaphors not only because of their ability to appreciate the point of metaphors or the similarities they are directed to perceive, but, more importantly, because they can reflect on, and understand, the content of metaphors. This observation is common to both literal and metaphorical uses of language. The crude causal account cannot adequately explain how hearers of metaphors can be become effective users of metaphors.

Another crucial feature of our practices of using metaphors is our capacity to use metaphors in arguments and engage in drawing certain inferences and implications from metaphorical sentences (Kwesi, 2018a; 2019a; 2019b). Consider these two arguments from Martinich (1996, p. 431, 435):

(1). "My love is a red rose.

A red rose is beautiful, or sweet smelling, or highly valued…

Therefore, my love is beautiful, or sweet smelling, or highly valued…"[14]

(2). "No man is an island

Every island is separated from every other thing of its own kind, does not depend upon any other thing of its own kind for its existence or well-being, and is not diminished by the destruction of any other of its own kind; …

Therefore, no man is separated from every other thing of its own kind, does not depend upon any other thing of its own kind for its existence or well-being, and is not diminished by the destruction of any other of its own kind" (Martinich, 1996, p. 435).[15]

Martinich considers (1) a valid argument and (2) an invalid argument. We need not worry about the validity of the arguments containing metaphors; it is enough to see that metaphorical sentences can serve as premises in arguments;

[14] The ellipsis is intended to show the open-endedness of the metaphor

[15] Martinich, however, thinks that "no man is an island" is not a metaphor. For according to him, "every metaphorical proposition is false" (1996, p. 430) and "it is true and not false that no man is an island" (p. 435) although he concedes that Donne's line is a figure of speech. I regard it as a metaphor because I do not subscribe to the view that the identifying mark of a metaphor is literal falsity.

metaphorical sentences play a role in reasoning. As premises of arguments, they can serve as reasons and justifications for conclusions, and stand in need of reasons and justifications (Kwesi, 2019a). The observation that metaphors can serve as premises and conclusions of arguments and that we can draw inferences from the metaphors we put forward suggests that metaphors must have meanings and contents. In reasoning with metaphors, users and their audiences are able to make inferences from the metaphors and provide other statements (metaphorical and literal) that tend to extend and explicate further the meaning and import of the metaphors. Consider the popular Psalm 23 from the Bible:

> The Lord is my shepherd, I shall not want.
>
> He maketh me to lie down in green pastures;
>
> He leadeth me beside the still waters.
>
> He restoreth my soul:
>
> He leadeth me in the paths of righteousness, for his name's sake.
>
> Yeah, though I walk through the valley of the shadow of death,
>
> I will fear no evil: for thou art with me,
>
> Thy rod and thy staff, they comfort me.

What is going on here in the psalm is that the Psalmist starts with the metaphor "the lord is my shepherd" and provides inferences that we can draw from the metaphor: if the lord is my shepherd then I shall not want, he will lead me to green pastures, his rod and staff with comfort me, etc.[16] Tirrell (1989) has called this phenomenon the "extending of metaphor". "The Lord is my shepherd" in our example is for her the "initiating metaphor" and the other expressions as the "extensions" of the metaphor which together with the initiating metaphor form a "metaphorical network or chain". The Psalmist presents us with an inferential metaphorical network where we see that his not wanting and being led to green pastures follow from his initial metaphor that the lord is his shepherd. For Tirrell, understanding a metaphor amounts to being "able to make appropriate uses of its extensions" (p. 18). Sometimes, the metaphor-maker herself provides the various extensions of the metaphor which develop and explain the metaphor in more detail. An example Tirrell uses is from Shakespeare's *Romeo and Juliet* where Lady Capulet not only tells Juliet to "read o'er the volume of Young Paris' face" but extends her metaphor to provide better and further particulars of what is involved in this kind of reading:

[16] This does not suggest that the Psalmist himself is actually making these inferences and connections; it is enough for his audience and readers of the Bible to draw these connections as they ponder on the initial metaphor.

Read o'er the volume of Young Paris' face,

And find delight writ there with beauty's pen;

Examine every married lineament,

And see how one another lends content;

And what obscur'd in this fair volume lies

Find written in the margent of his eyes.

This precious book of love, this unbound lover,

To beautify him, only lacks a cover. (Act 1 Sc 3)

Tirrell herself stops short of saying that the extensions of the metaphor serve as unpacking the meaning and content of the metaphor; her interest is in showing how extended metaphors impact our understanding of metaphors. However controversial the relation between the original metaphor and its extensions could be, the possibility of providing extensions to a metaphor suggests that the original metaphor had a meaning and a content, for it does not sound intuitive to suggest that the effects of the metaphor were being extended. The meaning of a metaphor can be extended; its effects cannot be analogously extended.

A final feature about metaphor worth noting is that we can agree and disagree with, assent and dissent to, certain metaphorical utterances. Such agreements and disagreements reflect our understanding of metaphors—we cannot agree or disagree on a metaphor if we do not understand it. Also, if we can agree or disagree over a metaphorical sentence then it implies that the sentence has been recognized or identified as a metaphor, since a genuine disagreement cannot obtain between two people over a particular sentence if one construes the sentence metaphorically and the other understands it literally. The two people clearly agree about the statement as put forward literally, but disagreement only emerges when the sentence is considered metaphorically. For instance, where Tom asserts that "the vice-chancellor is a bulldozer" and Harry responds by saying that "No, that's not true", the use of "that" here refers to the proposition expressed by Tom's assertion. Harry's response here expresses his disagreement with the content of the assertion made by Tom.

If, as argued by Davidson and Reimer (2001), a metaphorical assertion like "the vice chancellor is a bulldozer" only has a literal content or expresses a literal proposition, Harry's response will be conversationally infelicitous or inappropriate. For, the metaphorical assertion is literally false and hence, responding to it by saying that "that's not true" or "that's false" is both inappropriate and uninformative. But if Harry is warranted in making his response, if he is understood to be denying the assertion made by Tom, and if his use of "that" refers to the proposition expressed by Tom's assertion, then it is plausible to suppose that there is a propositional content other than the literal content of Tom's assertion that Harry rejects here. The intuitive conflict in the dialogue between Tom and

Harry can be attributed to the content expressed by Tom's assertion that Harry disagrees with. In our practices of using metaphors we can have disagreements—disagreements not merely over the significance or effects of metaphors but the contents expressed by the metaphors. The notion of disagreement primarily involves an incompatibility in the attitudes of the disagreeing parties towards a particular proposition. And if disagreements can occur with metaphors then we can infer that metaphors must have contents for disagreements to be possible.

5. CONCLUSION

The above criticisms of the denial of the meaning and content of metaphors and the arguments we adduced in favour of metaphors having contents, suggest the following desiderata for a satisfactory account of metaphorical content:

1. *Non-compositionality*: The account should explain how the content of a metaphor is not arrived at by a compositional analysis, although, unlike that of an idiom, the content is connected to the meanings of the constituents that make it up.
2. *Figurativeness*: The account should explain the ways in which the metaphorical is distinct from the literal in terms of the derivation of their contents.
3. *Disagreement*: The account should be compatible with how there can be genuine disagreements involving metaphors.
4. *Assertion and Retraction*: The account should show speakers' ability to put forward claims and stand by those claims or retract earlier claims. It should also be able to explain how metaphors can serve as premises and conclusions of arguments.
5. *Inference and Extension:* The account should explain speakers' ability to make inferences from metaphorical claims and be able to extend and explicate original metaphors
6. *Use in Soliloquies*: The account should make sense of speakers' use of metaphors in monologues and in soliloquies where there are no intended audiences.
7. *Hearers' Uptake*: The account should be able to explain hearers' immediate understanding of metaphors and their ability to use the metaphors in other contexts to produce effects on their hearers.

We can conclude that the Davidsonian causal account of metaphor gets it right by arguing that the words in a metaphor do not have additional or extra meanings other than their literal meanings. The account also seems plausible in indicating that metaphor also "nudges", "provokes", and "intimates" us to do things in certain ways because they have effects on us. But for reasons given

above, the account is not favourable in its further thesis that metaphors having effects is opposed to their having content, meaning or truth. I have tried to show that one can accept the central thesis of the causal account of metaphor (when it is understood in terms of word-literalism) and still posit that metaphors have content and meaning. The fact that we can use/misuse metaphors, that we can understand/misunderstand metaphors, that we can agree/disagree with metaphors, and the fact that we do reason with metaphors in arguments and make inferences from metaphors—all go to show that associated with a metaphor is a propositional content that we can grasp and evaluate.

REFERENCES

Beardsley, M. (1962). Metaphorical Twist. *Philosophy and Phenomenological Research*, *22*(3), 293–307.

Beardsley, M. (1978). Metaphorical Senses. *Noûs*, *12*(1), 3–16.

Black, M. (1979). How Metaphors Work: A Reply to Donald Davidson. In: S. Sacks (Ed). *On Metaphor* (pp. 181–192). Chicago: The University of Chicago Press.

Black, M. (1993). More About Metaphor. In A. Ortony (Ed.), *Metaphor and Thought.* (2nd ed., pp. 19–41). Cambridge: Cambridge University Press.

Camp, E. (2006a). Metaphor and that Certain "je ne sais quoi". *Philosophical Studies*, *129*(1), 1–25.

Camp, E. (2006b). Metaphor, Contextualism, and What Is Said. *Mind & Language*, *21*(3), 280–309.

Camp, E. (2006c). Metaphor in the Mind: The Cognition of Metaphor. *Philosophy Compass*, *1*(2), 154–170.

Camp, E. (2008). Showing, Telling, and Seeing: Metaphor and "Poetic" Language, *The Baltic International Yearbook of Cognition, Logic, and Communication: A Figure of Speech: Metaphor*, *3*, 1–24.

Carston, R. (1988). Implicature, Explicature and Truth-Theoretic Semantics. In: R. M. Kempson (Ed.), *Mental Representations: The Interface Between Language and Reality* (pp. 155–81). Cambridge: Cambridge University Press.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and Representations.* New York: Columbia University Press.

Cobreros, P. (2008). Supervaluationism and Logical Consequence: A Third Way. *Studia Logica*, *90*(3), 291–312.

Cohen, T. (1975). Figurative Speech and Figurative Acts. *Journal of Philosophy*, *72*(19), 669–684.
Cooper, D. (1986). Metaphor. New York: Basil Blackwell.
Davidson, D. (1967). Truth and Meaning. In: D. Davidson (Ed.), *Inquiries into Truth and Interpretation* (pp. 17–42). Oxford: Clarendon Press.
Davidson, D. (1979). What Metaphors Mean. In: S. Sacks (Ed), *On Metaphor* (pp. 29–45). Chicago: The University of Chicago Press.
Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
Davies, M. (1982). Idiom and Metaphor. *Proceedings of the Aristotelian Society, New Series*, *83*(1), 67–85.
Davies, S. (1984). Truth-Values and Metaphors. *The Journal of Aesthetics and Art Criticism*, *42*(3), 291–302.
Dever, J. (2008). Compositionality. In E. Lepore, B. Smith (Eds.), *The Oxford Handbook of Philosophy of Language.* Oxford: Oxford University Press.
Farrell, F. (1987). Metaphor and Davidsonian Theories of Meaning. *Canadian Journal of Philosophy*, *17*(3), 625–642.
Fine, K. (1975). Vagueness, Truth and Logic. *Synthese*, *30*(3/4), 265–300.
Fodor, J., Lepore, E. (2002). *The Compositionality Papers*. Oxford: Oxford University Press.
Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols.* Indianapolis: The Bobbs-Merrill Company Inc.
Goodman, N. (1979). Metaphor as Moonlighting. *Critical Inquiry*, *6*(1), 125–30.
Hall, A. (2009). Semantic Compositionality and Truth-Conditional Content. *Proceedings of the Aristotelian Society, New Series*, *109*(1pt3), 353–364.
Hesse, M. (1987). Unfamiliar Noises II: Tropical Talk: The Myth of the Literal. *Proceedings of the Aristotelian Society*, *61*, 283–311.
Hesse, M. (1988). The Cognitive Claims of Metaphor. *The Journal of Speculative Philosophy*, *2*(1), 1–16.
Johnson, M. (2008). Philosophy's Debt to Metaphor. In: R. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 39–52). Cambridge: Cambridge University Press.
Katz, J. (1973). Compositionality, Idiomaticity, and Lexical Substitution. In: S. R. Anderson, P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 357–376). New York: Holt, Rinehart & Winston.
Keefe, R. (2000). *Theories of Vagueness*. Cambridge: Cambridge University Press.
Keefe, R. (2008). Vagueness: Supervaluationism. *Philosophy Compass*, *3*(2), 315–324.
King, J., Stanley, J. (2005). Semantics, Pragmatics, and the Role of Semantic Content. In: Z. Szabo (Ed.), *Semantics Versus Pragmatics* (pp. 111–64). Oxford: Oxford University Press.
Kracht, M. (2011). Interpreted Languages and Compositionality. *Studies in Linguistics and Philosophy*, *89*, 1–223.

Kwesi, R. (2018a). Metaphor, Truth, and Representation. In P. Stalmaszczyk (Ed.), *Objects of Inquiry in Philosophy of Language and Linguistics* (pp. 117–146). Berlin: Peter Lang.

Kwesi, R. (2018b). Resemblance and Identity in Wallace Stevens' Conception of Metaphor. In K. Bartczak & J. Macha (Eds.). *Wallace Stevens: Poetry, Philosophy, and Figurative Language* (pp. 113–137). Berlin: Peter Lang.

Kwesi, R. (2019a). An Inferential Articulation of Metaphorical Assertions. *RIFL*, *13*(1), 116–132.

Kwesi, R. (2019b). Davidson's Phenomenological Argument Against the Cognitive Claims of Metaphor. *Axiomathes*, 1–24.

Lasersohn, P. (2012). Contextualism and Compositionality. *Linguistics and Philosophy*, *35*(2), 171–189.

Leddy, T. (1983). Davidson's Rejection of Metaphorical Meaning. *Philosophy & Rhetoric*, *16*(2), 63–78.

Lepore, E., Stone, M. (2010). Against Metaphorical Meaning. *Topoi*, *29*(2), 165–180.

Lepore, E., Stone, M. (2015). *Imagination and Convention: Distinguishing Grammar and Inference in Language.* Oxford: Oxford University Press.

Martinich, A. (1996). *The Philosophy of Language*. Oxford: Oxford University Press.

Moran, R. (1989). Seeing and Believing: Metaphor, Image and Force. *Critical Inquiry*, *16*(1), 87–112.

Nunberg, G., Sag, I. A., Wasow, T. (1994). Idioms. *Language*, *70*(3), 491–538.

Pagin, P., Westerstahl, D. (2010a). Compositionality I: Definitions and Variants. *Philosophy Compass*, *5*(3), 250–264.

Pagin, P., Westerstahl, D. (2010b). Compositionality II: Arguments and Problems. *Philosophy Compass*, *5*(3), 265–282.

Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press.

Reimer, M. (2001). Davidson on Metaphor. In: P. A. French, H. K. Wettstein (Eds.), *Midwest Studies in Philosophy, Volume 25: Figurative Language* (pp. 142–55). Boston, Mass, Oxford, UK: Blackwell.

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.

Rorty, R. (1987). Unfamiliar Noises I: Hesse and Davidson on Metaphor. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *61*, 283–311.

Rorty, R. (1989). *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.

Sperber, D., Wilson, D. (1995). *Relevance: Communication and Cognition.* Oxford: Blackwell.

Stanley, J. (2000). Context and Logical Form. *Linguistics and Philosophy*, *23*(4), 391–434.

Stanley, J. (2002). Making it Articulated. *Mind and Language*, *17*(1–2), 149–168.

Szabo, Z. (2010). The Determination of Content. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *148*(2), 253–272.

Tirrell, L. (1989). Extending: The Structure of Metaphor. *Noûs*, *23*(1), 17–34.
Titone, D. A., Connine, C. M. (1999). On the Compositional and Noncompositional Nature of Idiomatic Expressions. *Journal of Pragmatics*, *31*, 1655–1674.
Williamson, T. (1994). *Vagueness*. London: Routledge.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 135–153
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.08
Article

ALYONA BUDNIKOVA *

LEXICAL CONCEPTS AS FLUCTUATING STRUCTURES

SUMMARY: Lexical concepts (i.e. semantic units conventionally associated with linguistic forms) are viewed in the article as structures consisting of interrelated facets (i.e. conceptual slots filled with various types of information about the referent) with different structural weight. The paper suggests a way to model the graded structure of lexical concepts by assessing the weight of each constituting facet according to its relevance for defining purposes, frequency of contextual profiling and salience in derivation processes. Thus, the approach taken exploits as many linguistic points of access to the concept as possible and uses three different dimensions to range its facets. The suggested idea is verified with a case study of some common lexical concepts in English (e.g. represented by concrete nouns such as "bird", "tree", etc.), which reveals both the advantages and the limitations of the approach taken.

KEYWORDS: lexical concept, feature weights, ways-of-seeing, facets, meaning, profiling, salience.

INTRODUCTION

In cognitive linguistics, meaning is understood as not residing in the words as such, but pertaining to the level of actual expressions in speech, and the linguistic units are viewed rather as prompts for listeners to form meaningful conceptual representations (Radden, Köpcke, Berg, & Siemund, 2007, p. 1). As for words as such, their meanings are viewed as construed on-line on the basis of some flexible, open-ended semantic values (Evans, 2006, p. 491) with the help of contextual information and background knowledge (Lakoff, 1987; Langacker,

* Minsk State Linguistic University, Department of General Linguistics. E-mail: geronimo1981@mail.ru. ORCID: 0000-0002-4912-3234.

1987; Allwood, 1999; Croft & Cruse, 2004; Kharitonchik, 2009, etc.). The activation of the semantic value of a word in context, in its turn, allows the receiver (listener or reader) access to a certain facet of a richer cognitive structure associated with this lexical item (Langacker, 1991, p. 4; Taylor, 2006, p. 570). Thus, the multiple instances of word use in communicative situations "lead language users to decontextualize its meaning and to slowly grasp it as a gradually emerging entity which might change at any moment under the influence of a novel communicative situation" (Kharitonchik, 2009, p. 119). Viewed from this perspective, a lexical concept becomes "an inherently dynamic structure of semantic components" (p. 120) with one least changeable constituent—the classifier. The question for researchers, though, is how to objectively assess the centrality, or structural weight, of other components of a lexical concept which group around its classifier.

The question itself is not new. The idea of structuring the content plane of a lexical unit was exploited in structuralist semantics, where the components of lexical meaning were divided into linguistically relevant, necessary and sufficient to distinguish the given word from others in the lexical system (cf. the notion of "value of a linguistic sign" introduced by F. De Saussure), and extralinguistic components, also termed "potential", "occasional", etc. Thus, the semantic content of a word was looked upon in two aspects—the narrower and the wider—and one of the most important tasks (which is still a lexicographic problem) was to find objective criteria to delineate the linguistic from the encyclopaedic.

From the point of view of cognitive psychology, there are two major theoretical standpoints on this issue: the conceptual and the categorial (Khalidi, 1995; Sloman, Love, & Ahn, 1998). The conceptual centrality of a feature in the structure of any concept depends on whether it constitutes an essential property of the referent, determines its other properties, and whether its absence "affects the likelihood that an object belongs to a certain category" (Ahn, Kim, Lassaline, & Dennis, 2000, p. 361). For instance, in natural kind terms like "tiger," "swan," etc. the most conceptually central are internal, m o l e c u l a r features of the referents that cause all other characteristics, like appearance or behavior (Ahn et al., 2000, p. 362; Keil, Kim, & Grief, 2002, p. 380). In artefact terms like "chair," "table," etc. the conceptual core is constituted by functional features which determine all other properties of the referents (Sloman et al., 1998, p. 191; Keil et al., 2002, p. 380). However, this view does not explain why a complete ignorance of essential characteristics may not necessarily lead to a failure in reference. In its turn, the categorial centrality of a feature in the structure of a concept is determined by its vividness, distinguishing force, typicality, and diagnosticity (Pattabhiraman 1992, p. 27; Sloman et al., 1998, p. 192). It differs from the first perspective in reliance on external rather than internal properties when forming a class of entities.

However, from both standpoints the criteria of centrality largely depend on the type of mental operation—careful analysis of all features vs. brisk judgement

about categorial membership—and, consequently, are not based on the objective data of language use.

This paper continues the search for the principles of objective structuring of the content plane of a word and suggests using linguistic data as a marker of the structural weight of semantic features, thus, unifying the structuralist and cognitive approaches to lexical semantics.

METHODOLOGY

As mentioned in the introduction, the meaning of a word is a property of situational language use, and under various contextual factors we may observe the foregrounding of some aspects (also termed "facets", "ways-of-seeing" [Croft & Cruse, 2004, p. 137], etc.) of the designated concept and suppression of others. The well-known examples include cases with artifact terms like the word "book" which allows for the [TOME] and the [TEXT] reading in different contexts (Croft & Cruse, 2004, p. 116):

(1) Some of the books were **paperback**, most were **hardbacks**.

(2) Some of the books we read were novels and the others were **biographies**.

Whereas these readings, or facets, seem quite autonomous due to the metonymic link between them, in other contexts the activated facets seem more closely connected to each other, for example, the locative and behavioral facets in the concept BIRD in contexts (3) and (4):

(3) Recovering my composure, if not much dignity, I **followed** the bird **through the trees** [CRJ 1407][1] ('movement');

(4) Fears […] have been partly eased by preliminary studies of bird **habits** [AAL 501] ('typical actions').

These readings are not the ones usually pointed out in the lexicographic definitions of the word "bird", and the designated class of entities is the same in both examples (although the referents are different). However, the contextual focus is obviously on different aspects of the concept BIRD.

In fact, the observed similarity in the contextual modulations of noun concepts has led some scholars to introduce general lists of semantic aspects (facets, roles) which govern the linguistic manifestation of nouns. For instance, J. Pusteyovsky suggests four roles in the so-called *qualia structure* of nouns: the constitutive role (the internal constitution of the object, material, weight, parts, etc.); the formal role (the features distinguishing the object from other objects in

[1] The examples are taken from the British National Corpus (2007) and cited with the corresponding code.

a larger domain, such as shape, orientation, colour, etc.); the telic role (the function of the object); and the agentive role (how the object was created) (Pusteyovsky, 1991, p. 426–427).

The ways-of-seeing (WOS) introduced by Croft & Cruse include the part-whole WOS (views an entity as a whole with parts), the kind WOS (views an entity as a kind among other kinds), the functional WOS (views an entity in terms of its interaction with other entities) and the life-history WOS (views an entity in terms of its coming into being; Croft & Cruse, 2004, p. 137).

All the above-mentioned is also in line with the idea expressed by A.Wierzbicka, that

> the very complex structures of concrete concepts have a simple skeleton, a skeleton shared by many, if not all the concepts of a given domain, so that all the concepts of a given domain can be viewed as different answers to the same basic conceptual questionnaire. (Wierzbicka, 1985, p. 332)

The introduction of these general schemata has a rich explanatory potential when it comes to understanding how concepts combine in complex linguistic expressions (Murphy, 2002, p. 453), however, it does not show w h i c h o f t h e f a c e t s a r e m o r e s a l i e n t. Besides, without substantial empirical support, it is difficult to say whether the proposed lists of facets are finite.

The starting point in the search for the objective empirical criteria of evaluating feature weights in a concept is to consider the frequency parameter which, according to Dirk Geeraerts, underlies the overwhelming salience phenomena in lexicon and may be looked upon from onomasiological, semasiological and structural perspectives (Geeraerts, 2006, p. 74–94). Since the present research is semasiological in nature, let us consider the last two approaches.

Semasiological salience is "a relationship among various semantic possibilities of a given lexical item" (Geeraerts, 2006, p. 79), i.e. some semantic possibilities are chosen more often than others. It may be revealed paradigmatically (on the systemic level) and syntagmatically (in the language use), cf. "type frequency" vs. "token frequency" of a linguistic phenomenon (Bybee, 2003, p. 11–12). For example, the more frequently a certain reading is used with the linguistic unit, the more salient it is among other possible readings of the unit. And the more frequently the unit is used in a certain syntagmatic context, the more salient this context is among other possible syntagmatic contexts of the unit.

If we continue this theoretical line, the facets of a lexical concept salient paradigmatically are the ones that are most frequently used in association[2] with the concept, for example, when one tries to define it. This can be revealed not only by experimental (psycholinguistic) methods (Rosch, 1978, p. 32), but also with the help of lexicographic sources. For example, explaining the meaning of the

[2] The paper does not include free associations that spring to mind in relation to a given word, since they may not be conventional. However, if they are fixated in phraseology, they become part of the modelled structure.

word *bird* one may rely on perceptual and locative features of the referent (5); others will add biological properties to the list (6); still others will profile perceptual, biological and behavioral types of knowledge (7):

(5) bird "a creature with **feathers** and **wings**, usually **able to fly**" (CALD 2003);

(6) bird "a creature with **feathers** and **wings**. **Female** birds **lay eggs**. Most birds **can fly**" (Collins COBUILD 2007);

(7) bird "an animal covered in **feathers**, with two **wings** for **flying** and a **hard pointed mouth** called a beak or a bill. Birds **build nests**, in which **female** birds **lay eggs**" (Macmillan).

So, the frequency of lexicographic profiling of certain facets in the structure of lexical concepts will show how salient the facets are from the paradigmatic perspective. Likewise, the frequency of contextual manifestation of certain facets will testify to their syntagmatic salience.

Structural salience, after D. Geeraerts, is determined by the frequency with which a feature occurs "in the structure of the lexicon", i.e. in the "totality of distinctive relations in the lexicon" (Geeraerts, 2006, p. 88). One of the possible implementations of this dimension is to consider what features of a source concept occur as distinctive ones among the derived units from the given word. Consider some of the expressions derived from the word "bird" or using the concept BIRD as a starting point to form new linguistic expressions (8), (9), (10), (11), (12):

(8) bird 2: "a light object shaped like a cone that is hit over the net in the game of badminton; shuttlecock" (MWD);

(9) bird of passage: "a person who moves from place to place frequently" (AHDEL);

(10) bird-voiced tree frog: "a species of frog in the Hylidae family easily distinguishable during the spring and summer when it gives its characteristic call" (Bird-Voiced Tree Frog, n.d.);

(11) eat like a bird: "to eat sparingly" (RHD);

(12) free as a bird: "at liberty, without obligations" (AHDI).

These derivatives, due to the metaphorical shift which underlies their formation, highlight some features of the source concept: feathers, moving in the air

(8); migration (9); ability to sing (10); nutrition habits[3] (11); behavior (12). From this perspective the weight of a feature in a lexical concept is graded according to the frequency with which the feature was used as a base for a metaphorical shift.

The three types of linguistic manifestation of the facets constituting a lexical concept described above will constitute a basis to assess their structural weight. Thus, in this research the problem of grading the facets of a lexical concept will be solved using the established methodology of linguistic analysis.

RESEARCH

The case study was based on some common English concrete nouns (12 animal and plant terms, 7 terms referring to people) characterized by the diversity of the constituents of the corresponding lexical concepts. The choice of lexical items was mostly based on their relatively high contextual frequency[4] and difference in the levels of categorization among them (basic level, superordinate level, and subordinate level terms).

The sources of information about the features of the analyzed lexical concepts included their lexicographic definitions in 10 English-English dictionaries (190 entries used), corpus data of the profiled features in more than 2000 contexts for each word retrieved from the British National Corpus (2007), and the derived units foregrounding certain properties of the source concepts (358 derived linguistic expressions) found in idiom dictionaries and other etymological sources.

The choice of dictionaries was based on the premise not to miss any aspect or type of information that is relevant for definitional aims, that is why the list included American dictionaries alongside British ones, learner's dictionaries together with unabridged lexicographic sources, older and contemporary dictionaries (see the full list in the cited literature). The facets and the features constituting them were discovered in the process of componential analysis of the studied definitions. To unify all the classifiers, we substituted the narrower ones with their definitions. For example, "mammal" was substituted with "a warm-blooded vertebrate animal of a class that is distinguished […]", so that to explicate a broader classifier "animal", common for all the studied animal terms. As a result, some definitions were enriched with new features coming from transformations of this kind.

The range of contextual word combinations under analysis included four basic structural types: noun phrases and predicative phrases with the studied words as heads (e.g. "The **lean-flanked wild** and **free** horse […]" [A0L 3852]; "[…] the cat **curled up** against his feet" [FPB 618]), noun phrases with the stud-

[3] Phraseological units fixate different features of the source concept, be it scientifically proved properties or folk wisdom. As a result, the image of the referent in this perspective may be far from reality, like the sparse nutritional habits of birds.

[4] Most of the analysed words belong to the 3000 most frequent words from Longman Communication 3000 (2nd ed.).

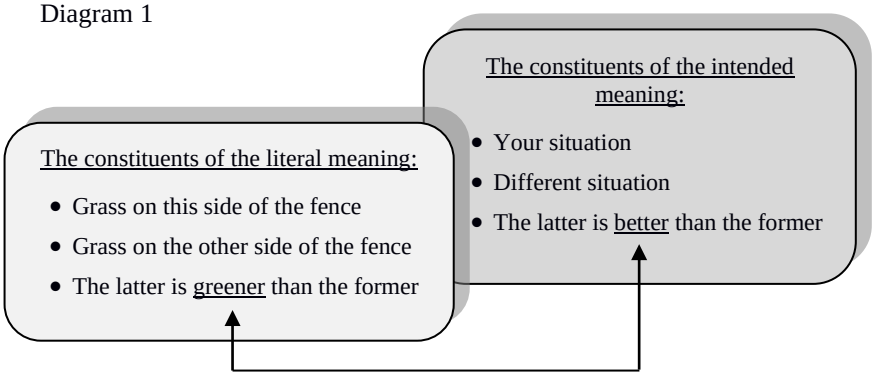
ied words as modifiers (e.g. "[…] bird **droppings**" [ADA 1675]; "[…] the normal **life span of** the mouse" [EA0 1103]), verb phrases with the studied words as arguments (e.g. "[…] remember to **pat** the dog" [CJE 553]; […] "swooping in **from behind** the owl" [BLX 1644]), and phrases with subordinate clauses attributive to the studied words (e.g. "The wave is a fish **that always gets away**" [ASV 82]; "[…] the creature **which has never been seen before in any swamp or tropical rainforest**" [AKE 155]). The contexts not taken into consideration were scarce and included the following main types: 1) the referent of the studied word belongs to a different category, e.g. "[…] **artificial** grass" [CJC 533] (not a category of plants); 2) the studied word refers to the class in general, without specifying any aspect, e.g. "[…] to **learn more about** the fish" [J3N 147], "Donald Gillies **was** a friend" [A0N 1754]; 3) the studied word is a component of a set nominative complex, e.g. "[…] plant **life**" [B0P 556].

The analysis of the derived and phraseological units under consideration consisted of three stages: 1) stating the literal and intended meaning of a derivative, 2) identifying the constituents of both that are being mapped, and 3) finding the semantic links between them (Lakoff & Turner, 1989, p. 170–191; Geeraerts, 2006, p. 200). Consider example (13):

(13) The grass is always greener on the other side (of the fence)**:** "a different situation always seems better than one's own" (AHDI).

The intended meaning of the phrase is made possible due to the mapping of the colour characteristics of the source concept GRASS and evaluative characteristics of the target concept SITUATION (see Picture 1). Consequently, the more colour there is in the grass, the higher the value of the situation is. Therefore, the perceptual feature "colour" of the concept GRASS is made prominent via this set expression.

Diagram 1



Another example is (14) which highlights the biological and temporal facets of the concept GRASS ("low speed of growth"—"delays in acting"):

(14) not let the grass grow under one's feet: "not delay in acting or taking an opportunity" (OD).

Conversely, in metonymy based expressions like grassy: "covered with grass", grass box: "a container attached to a lawn mower that receives grass after it has been cut", grass widow:[5] "a woman who spends a lot of time apart from her partner, often because he or she is working in a different place" (CALD), etc. we cannot identify the features of the source concept that get fixated in the derived concepts. So, suchlike expressions were not taken into account.

The frequencies of feature profiling in dictionary definitions were calculated as percentages of the number of features found to the number of definitions (10 definitions for each word). The contextual frequency is a relation of the number of profiled features to the 2000 sentences found for each word. Finally, the frequency in derivation processes was estimated as the number of features found as bases for metaphors to the whole number of the derivatives (it is specified for each of the studied words and ranges from 56 for "dog" to 1 for "oak" and "person").

All the obtained frequencies were later unified to a 10-grade scale using the formula in (Baltatescu, 2002). This allowed the building of models of lexical concepts reflecting the different structural weights of their constituting facets, or types of features.

The research shed light on some important aspects of lexical semantics which I would like to dwell upon in detail in the next section.

FINDINGS AND DISCUSSION

1. The Open-Endedness and Relative Stability of Lexical Concepts

The first thing clearly observed from each of the studied perspectives (definitional needs, contextual profiling, and derivational activity) is the following: no matter how big the variety of the profiled information about the referents of the words is, it is logically reducible to a closed set of conceptual slots. There are 10 facets for animal terms, 8 for plant terms, 9 for terms referring to people. However, the list is longer than the one suggested by J. Pusteyovsky and Croft & Cruse, which is explained by the broader scope of the employed empirical data.

Consider the features of the lexical concept BIRD foregrounded in its various types of linguistic manifestation (see Table 1).

[5] According to etymological sources, the allusion to grass is not clear, but it was commonly believed to refer to casual bedding (see Online Etymology Dictionary, 2019).

Table 1

The features of the lexical concept BIRD obtained from linguistic data

Semantic features and facets in the lexical concept BIRD	Profiled		
	in definitions	in context	via derivatives
Perceptual (size, appearance, shape, etc.)	+	+	+
Systematizing (taxonomic name, etc.)	+	+	
Biological (biological activity, gender, etc.)	+	+	+
Constitutive (internal parts, container, etc.)	+	+	+
Behavioral (habits, situational actions, etc.)	+	+	+
Locative (movement, habitat, location, etc.)	+	+	+
Utilitarian (domesticity, owner, etc.)		+	
Temporal (age, stage of development, etc.)		+	+
Psychological (character, emotional state, etc.)		+	+
Social (role in society, pedigree, etc.)		+	

It is clearly observed that some of the discovered facets of the concept BIRD are specific to the domain of animals and animate objects in general (biological, behavioral, psychological features); others reveal the connection of the designated concept with a broader concept THING (perceptual, constitutive, locative, temporal features). Also, the list includes the so-called meta-facet of systematizing features (i.e. relevant in human-built taxonomies) and the utilitarian facet showing the proximity of the referent to humans. The same refers to all other concepts under study: the majority of the facets serve as general links to other concepts, and there are certain anthropocentric facets.

As the data show, although the contextual use is clearly the richest source of features of the analyzed lexical concepts, the number of rubrics that the information obtained can be divided into does seem to be rather stable, and new features that might be actualized under novel contextual circumstances are likely to fall under one of the categories already established for the first 1000 contextual findings. This demonstrates the stability and open-endedness of a lexical concept. The stability is achieved by the set number of facets that can potentially be present in lexical concepts belonging in the same broader domain (e.g. of

animate things). The open-endedness is achieved by the fact that within the established facets more and more features may appear under novel contextual factors. For example, in contexts (15), (16) we may observe the activation of the feature "container" not common for concepts represented by animate things. However, the new feature appears under the already established rubric (constitutive facet).

(15) […] the mercury compounds underwent further "biotransformation" **inside** the fish [AMS 576];

(16) […] a condom prevents sperm from **entering** the woman [A0J 447].

Finally, as the data show, the conceptual slots may be filled in one perspective and left empty in another (see Table 2).

Table 2

The number of the profiled facets of the lexical concepts

Some of the analyzed lexical concepts	The number of facets profiled		
	in definitions	in context	via derivatives
CREATURE	4	10	0
BIRD	6	10	7
CAT	6	10	7
HORSE	8	10	8
OWL	7	10	5
PLANT	5	8	0
GRASS	5	7	5
PERSON	5	8	1
WOMAN	3	8	5
TEACHER	2	8	2

This refers, first, to the facets actualized in one dimension and nonexistent in another. For example, the dimension of derivation serves as a good basis for evaluation feature weights only for concepts represented by basic and subordinate level terms. As for superordinate level terms, the data here are rather scarce, with only one derivative found: person 2: "biol. a shoot or bud of a plant; a polyp or zooid of the compound Hydrozoa, Anthozoa, etc." (Biology Dictionary, 2019), which profiles the constitutive facet of the source concept PERSON ("member of a group"—"part of the whole"). Secondly, the facets profiled in the three dimensions do not always overlap, which proves that a lexical concept cannot

be fully revealed in only one of its linguistic manifestations.

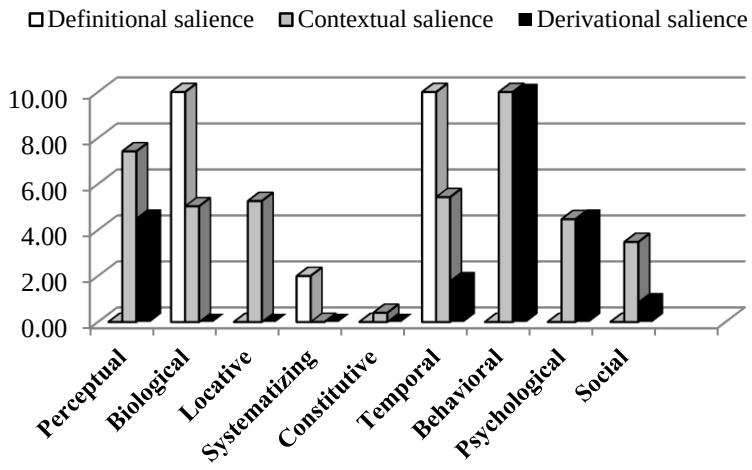
2. The Fluctuation of Lexical Concepts

In most of the studied concepts the rigid dividing line between central and peripheral constituents is impossible to draw, since, firstly, the weights of facets in different linguistic perspectives may not coincide, and, secondly, the difference in salience among them may not be great.

Consider the structure of the lexical concepts WOMAN and TEACHER. In the structure of the concept WOMAN (see Diagram 1) there are no facets equally salient in all three perspectives.

Figure 1

The model of the lexical concept WOMAN showing the weights of its facets



For example, the information relating to the behavior and character of the word's referents is most frequently profiled in context, as in (17), (18) or (19), and via derived units, as in (20) or (21), but is completely irrelevant for definitional needs (22):

(17) […] traditional female tasks such as cooking and cleaning are more likely to be shared equally when the <u>woman</u> **works full-time** [AP5 243] ('occupation');

(18) The woman **showed her a bundle of tiny packages, secret notes**, from some of the disappeared men, smuggled out of their secret jail by one of the guards [A03 832] (‘situational action’);

(19) In their defence, I’m not **the most approachable** woman in the world [A0F 506] (‘habitual behavior’);

(20) woman’s tongue: “bot. the tree *Albizia lebbeck*; an allusion to the sound the seeds make as they rattle inside the pots, and women’s tendency to gossip” (‘typical behavior’; Woman’s Tongue Tree, n.d.);

(21) womanish: “suggestive of a weak character” (MWD) (‘character’);

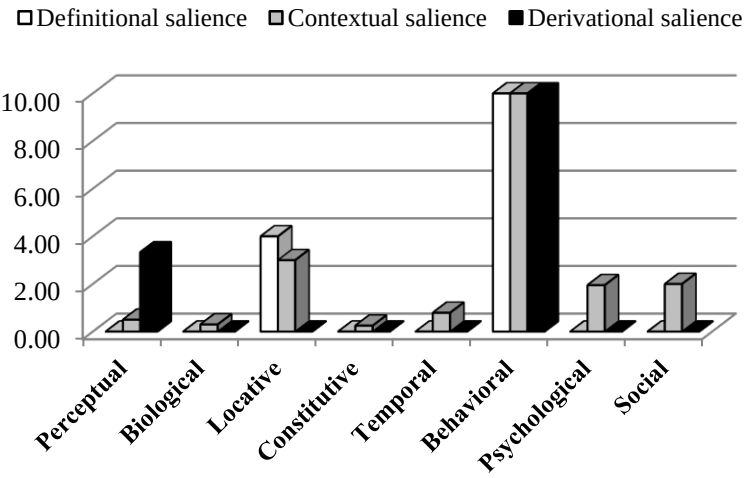
(22) woman: “an adult female human being” (LDOCE).

The same difference in structural weight is observed for biological and temporal facets of the lexical concept WOMAN which are salient for definitional needs, but have relatively low salience in context, let alone derivation.

In the lexical concept TEACHER (see Diagram 2), quite the opposite, there is a facet salient in all the perspectives—the behavioral one.

Figure 2

The model of the lexical concept TEACHER showing the weights of its facets



Behavioral features and, namely, the type of occupation of the word’s potential referents, are profiled in all its definitions, e.g. teacher: “someone **whose job is to teach**” (Macmillan Online). In contextual use the features relating to all

types of behavior of the referents are also most frequent, as in (23), (24), and (25):

(23) The case of a **physics** teacher [A06 189] (‘job specialization’);

(24) In such a setting, he believes, his **work as** a teacher can only be betrayed [A05 1599] (‘occupation’);

(25) Method must be learnt from a **specially-trained** teacher [A0J 474] (‘qualification’).

As for the sparse derivatives of the word “teacher”, they are all based on the features of the source concept in this or that way related to the referent’s occupation, as in (26) and (27):

(26) teacher 2: “the index finger; the forefinger” (“typical gesture—pointing”; Wordnik, 2019);

(27) teacher’s nodes: “small, circumscribed, bilateral, beadlike enlargements on the vocal cords caused by overuse or abuse of the voice” (“habitual behavior—speaking”; Mondofacto, 2019).

So, the behavioral facet in the structure of the concept TEACHER could be considered central in all the studied perspectives. However, other facets, being surely on the periphery, show a certain degree of fluctuation in salience according to the type of manifestation and, thus, are difficult to range objectively in one dimension.

To sum up, the structure of a concrete concept is multidimensional, i.e. in each type of its linguistic manifestation we observe different order of its semantic constituents.

The analysis of patterns of meaning fluctuation in the semantics of the studied words revealed both similarities and differences, and this makes it interesting to analyze the contributing factors.

3. The Difference in Patterns of Fluctuation

In the distribution of feature weights in the studied concepts there can be observed two tendencies: to the convergence and the divergence of most salient facets in all three perspectives. There are no clear-cut cases to illustrate, but the studied models of lexical concepts may be considered as being closer to this or that trend.

3.1. The Convergence Tendency

The tendency to the convergence, or overlap, of most salient facets in their different linguistic realizations is predominant in the studied material and can be

observed in the lexical concepts TEACHER and FRIEND, and also in the cases of BIRD, OWL, TREE and GRASS. The concepts HUSBAND and WIFE occupy an intermediary position, with only partial overlap of most salient facets across the three perspectives.

In the TEACHER and FRIEND cases the convergence tendency is most vivid, with one facet salient in all the perspectives and other components lagging far behind. The nouns "teacher" and "friend", being nominal kind terms, along with "husband" and "wife", differ from the latter in their functional character. Thus, even though their referents are human beings with biological, perceptual, locative, etc. properties that can be attributed to them, it's their function in society, emphasized in their definitions, that becomes the crucial factor guiding the linguistic realization of these concepts. For example, the above discussed contexts (23), (24), and (25) for "teacher", with the foregrounded behavioral facet, are more frequent than, say, (28) or (29), with the locative and perceptual properties profiled, though they are fairly possible:

(28) I was just sitting reading and the teacher **walked in** and I didn't hear him [KDP 2814] ('movement');

(29) I have met a very **nice young lady** teacher [A89 417] ('appearance', 'age', 'gender').

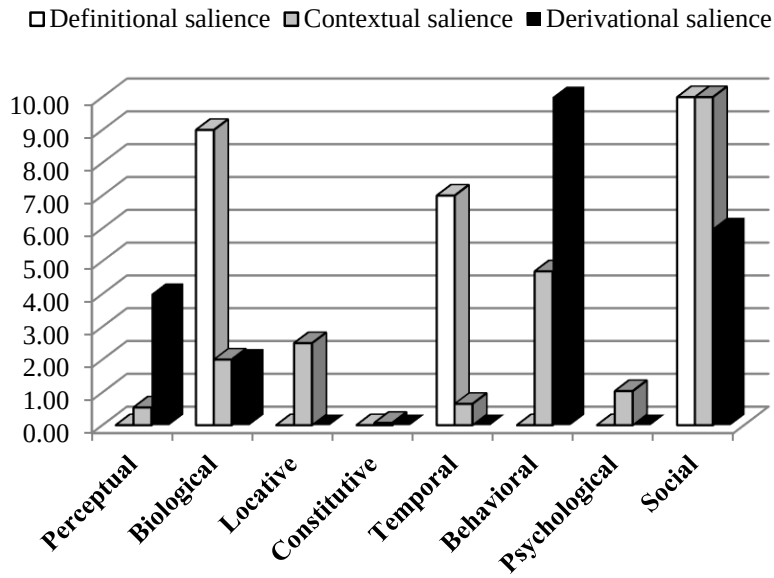
As for the derived units of different kinds built on the basis of the words "teacher" and "friend", they are also formed by exploiting the functional side (i.e. behavioral features) of the source concepts, with other features present only scarcely, accompanying the behavioral ones. In the example (26) there is a metonymic shift of *totum pro parte* type, where the term for the whole concept (a teacher with the highlighted typical gesture of pointing with a forefinger) is used to name part of the concept (the forefinger itself). At the same time, the source referent (a teacher) and a target referent (a human's forefinger in general) share a common functional characteristic (pointing), so the case might be treated as a metaphorical extension as well. At any rate, the feature foregrounded in the concept TEACHER by this derivative is "gesture" which is, on the one hand, a type of behavior, and, on the other hand, involves information about visible body parts (perceptual facet).

The lexical concepts HUSBAND and WIFE do not show the convergence tendency so vividly, but they are similar to TEACHER and FRIEND in the partial overlap of feature salience in all three perspectives. Namely, the social facet is most salient in definitions and context, but comes second in derivation (see Diagram 3).

The social facet is presented mainly by the feature "marriage partner", as in contexts like "my husband"; "her husband"; "the husband of the cleaning woman", etc. Thus, it's the relational character of the word "husband" (and also "wife") that is the key factor in structuring their meaning constituents in terms of salience for definitional needs and context profiling.

Figure 3

The model of the lexical concept HUSBAND showing the weights of its facets



As for the derivatives of the word "husband", they are mainly formed in reliance on typical behavior attributed to husbands (30) and some perceptual properties used to be characteristic of them in the past (31):

(30) husband: "to use something carefully so that you do not use all of it" (CALD 2003) ('typical behavior—economizing');

(31) husband 2: "a polled tree; a pollard; so called in humorous allusion to the traditional bald head of husbands with energetic wives" (CD) ('typical haircut—cropped').

In general, since the division into facets is not completely devoid of subjectivity and the social facet in relation to human referents might be considered similar to the behavioral one, the overall picture with the most salient constituent seems much the same as in the TEACHER case. The only difference, then, is the greater degree of fluctuation in salience of other facets in the meaning structure of "husband" compared to "teacher".

In the cases of BIRD, OWL, TREE, and GRASS, the facet most salient in all the dimensions is the perceptual one. However, other types of knowledge about the referents of these animal and plant terms (mainly, biological, locative and

utilitarian) do not lag far behind in salience. So, depending on the nature of the designated referent (its typical behavior, habitat and proximity to people) the distribution of structural weight among these dominant facets will differ, but not significantly.

3.2. The Divergence Tendency

The divergence tendency in salience, i.e. the absence of facets similarly high in frequency in all types of linguistic realization, is observed only in the lexical concepts MAN and WOMAN. The words "man" and "woman" are predominantly aimed at characterizing rather than identifying their referents, hence the relatively limited and stable number of features in their definitions (mainly "gender" and "age"). However, their actual use in context and as derivational bases triggers a vast range of otherwise latent types of knowledge constituting these lexical concepts (behavior, appearance, locomotion, character, social role, etc.). The difference in fluctuation patterns in the lexical concepts, represented by the nouns *man* and *woman*, and the concepts represented by other terms of people, also characterizing in their primary function, may be attributed to the features underlying the corresponding classes of referents: the so-called natural characteristics (gender and age) vs. the nominal characteristics (profession, marital status, interpersonal relations).

CONCLUSIONS

The suggested models of lexical concepts built on empirical data are more embracing than those built on the basis of psycholinguistic and introspective studies, since they include not only the types of knowledge relevant for definitional aims. At the same time, they do not go too far from the schemata proposed deductively, which serves as a proof of their objectivity.

The models of lexical concepts construed in the research demonstrate how these inherently stable structures may change and enrich themselves under novel circumstances. The stability of a lexical concept is achieved by the set number of constituting facets, common for the whole class of similar entities, but depending on the type of linguistic manifestation, the facets are either filled, or left empty; either frequent, or scarce. The enrichment of a concept with new features is possible within the established rubrics (facets); otherwise a new concept is formed.

The undertaken study illustrates that although in many cases the most salient facets in all the studied dimensions are the same, the weight of other facets is mostly fluctuating depending on the type of linguistic manifestation. The patterns of fluctuation of the facets in the studied concrete concepts differ due to the influence of the following main factors: the nature of the referent, the function that the corresponding word is primarily used in (identifying vs. characterizing), the features underlying the concept (natural vs. nominal), and the relational or functional character of the word content. The rigid dividing line between the

most salient facets and the rest is only possible to draw for the concept TEACHER, which is due to its highly functional character (i.e. the word "teacher" mainly serves to characterize the referent by its main function).

Among the limitations of the approach taken I would mention the derivation criterion which works for some concepts, but is useless for others (e.g. most superordinate terms) and the part-of-speech dependency (it works well only for multi-faceted concepts, like the ones represented by concrete nouns or some types of verbs).

ACKNOWLEGMENTS

I would like to thank my teacher, Prof. Z.A. Kharitonchik, PhD, for inspiration, valuable comments and support in my research.

REFERENCES

Ahn, W., Kim, N. S., Lassaline M. E., Dennis, M. J. (2000). Causal Status as a Determinant of Feature Centrality. *Cognitive Psychology*, *41*(4), 361–416.

Allwood, J. (1999). Semantics as Meaning Determination with Semantic-Epistemic Operations. In: J. Allwood, P. Gärdenfors (Eds.), *Cognitive Semantics. Meaning and Cognition* (pp. 1–18). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Baltatescu, S. (2002). *Problems of Transforming Scales of Life Satisfaction* [PowerPoint presentation]. Retrieved from: https://studylib.net/doc/9985031/problems-of-transforming-scales-of-life-satisfaction

Biology Dictionary. (2019). *Person*. Retrieved from: https://www.biology-online.org/dictionary/Person

Bird-Voiced Tree Frog. (n.d.). Retrieved from: https://en.wikipedia.org/wiki/Bird-voiced_tree_frog

Bybee, J. (2003). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Croft, W., Cruse, A. D. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Evans, V. (2006). Lexical Concepts, Cognitive Models and Meaning-Construction. *Cognitive Linguistics*, *17*(4), 491–534.

Geeraerts, D. (2006). *Words and Other Wonders. Papers on Lexical and Other Topics*. Berlin, New York: Mouton de Gruyter.

Keil, F., Kim N. S., Grief M. L. (2002). Categories and Levels of Information. In: M. E. Forde, G. Humphreys (Eds.), *Category Specificity in Brain and Mind* (pp. 375–401). Hove: Psychology Press.

Khalidi, M. A. (1995). Two Concepts of Concept. *Mind and Language*, *10*(4), 402–422.

Kharitonchik, Z. A. (2009). Lexical Meaning as a Dynamic Entity of Variant Semantic Components. *Probleme Actuale de Lingvistică, Glotodidactică şi Ştinţă Literară*, *IV*(I), 113–121.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal About the Mind*. Chicago, London: University of Chicago Press.

Lakoff, G., Turner, M. L. (1989). *More Than Cool Reason. A Field Guide to Poetic Metaphor*. Chicago, London: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Descriptive Application*. Stanford: Stanford University Press.

Mondofacto. (2019). Retrieved from: http://www.mondofacto.com/facts/dictionary?query=teacher%27s+nodes&action=look+it+up

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge (Massachusetts), London: The MIT Press.

Online Etymology Dictionary. (2019). Retrieved from: www.etymonline.com/

Pattabhiraman, T. (1992). *Aspects of Salience in Natural Language Generation*. (Ph.D. dissertation, Simon Fraser University, Vancouver).

Pustejovsky, J. (1991). The Generative Lexicon. *Computational Linguistics*, *17*(4), 409–441.

Radden, G., Köpcke, K.-M., Berg, T., Siemund, P. (2007). The Construction of Meaning in Language. In: G. Radden, K.-M. Köpcke, T. Berg, P. Siemund (Eds.), *Aspects of Meaning Construction* (pp. 1–15). Amsterdam (Philadelphia): John Benjamins Publishing Company.

Rosch, E. (1978). Principles of Categorization. In: E. Rosch, B. B. Lloyd (Ed.), *Cognition and Categorization* (pp. 27–48). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Sloman, S., Love, B., Ahn, W. (1998). Feature Centrality and Conceptual Coherence. *Cognitive Science*, *22*(2), 189–228.

Taylor, J. R. (2006). Cognitive Linguistics and Cognitive Semantics. In: K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (Vol. 2, pp. 569–582). Amsterdam: Elsevier Science.

Wierzbicka, A. (1985). *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma Publishers.

Woman's Tongue Tree. (n.d.). Retrieved from: https://en.wiktionary.org/wiki/woman%27s_tongue_tree

Wordnik. (2019). *Teacher*. Retrieved from: https://www.wordnik.com/words/teacher

Sources of Lexicographic Definitions:

AHDEL: *American Heritage Dictionary of the English Language*. Retrieved from: https://www.ahdictionary.com

CALD: *Cambridge Advanced Learner's Dictionary*. (2003). Cambridge University Press.

CED: *Collins English Dictionary. Complete and Unabridged*. Retrieved from: http://dictionary.reference.com
Collins COBUILD: *The Collins COBUILD Advanced Dictionary of American English*. (2007). Harper Collins Publishers Ltd.
LDOCE: *Longman Dictionary of Contemporary English*. Retrieved from: http://www.ldoceonline.com
Macmillan: *Macmillan Dictionary*. Springer Nature Ltd. Retrieved from: http://www.macmillandictionary.com
MWD: *Merriam-Webster Dictionary*. Retrieved from: http://www.merriam-webster.com
OD: *Oxford Dictionaries*. Retrieved from: http://www.oxforddictionaries.com
RHD: *Random House Dictionary*. Retrieved from: dictionary.reference.com/
WRUD: *Webster's Revised Unabridged Dictionary*. DICT Development Group. Retrieved from: http://www.dict.org/

Cited Sources of Idioms:

CD: *The Century Dictionary: An Encyclopedic Lexicon of the English Language*. Retrieved from: www.micmap.org/dicfro/introduction/century-dictionary
AHDI: *The American Heritage Dictionary of Idioms*. (1997). Houghton Mifflin.

Source of Contexts:

The British National Corpus, version 3. (2007). Oxford: Oxford University Computing Services, on behalf of the BNC Consortium.

STUDIA SEMIOTYCZNE, t. XXXIII, nr 1 (2019), s. 155–177
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/sts.xxxiii1.09
Article

WIESŁAW WALENTUKIEWICZ *

BASIC CONCEPTS: A COGNITIVE APPROACH

SUMMARY: This article seeks to describe concepts of a special kind, these being ones that count as basic, while at the same time referring to the results of research in logic, the philosophy of language, and empirically pursued cognitive psychology. The key issue addressed is this: on what grounds are such basic concepts formed? It thus investigates issues pertaining to their formation and operation, especially in small children. (Basic-level concepts will be examples of basic concepts.) Such concepts can take the form of mental representations of objects, properties and relations. They function in classifications made by numerous and diverse cultural groups, are established at an early stage—being the first to be named and, so to speak, malleable—and their structure is not satisfactorily captured by any currently recognized theory. Moreover, they are organized around some sort of overall similarity irreducible to any particular component part. (Basic concepts pertaining to properties and relations must be based on some overall similarity, as properties and relations themselves do not consist of parts. Equally, basic concepts pertaining to objects cannot be constructed on the basis of mere parts of these objects.) Psychologists and philosophers, on the other hand, frequently claim that properties are component parts to which overall similarity can be reduced (e.g. in exemplar-based and prototype-based theories of concepts). Yet if this solution were to be accepted, one would then have to say that three- or four-month-old children are unable to establish properties before delimiting the range of the relevant category (or any fragment of this range), whilst also being unable to establish the range of that category (or any fragment of it) before delimiting its properties. The problem with this is that children can distinguish some properties; however, they are incapable of establishing within a relatively short period of time which of these properties determine membership in the sense of falling within the range of the category in question. Moreover, basic concepts cannot be organized on the basis of a relation of similarity reducible to properties, due to the fact that any such similarity will be an equivalence relation, whilst the similarity relation accessible to the child constitutes a non-equivalence relation. A further point is that no consensus has yet been agreed upon within the psychological literature as to the construction of concepts formed by three- or four-month-old children.

* University of Economics in Katowice, Department of Communication Design and Analysis. E-mail: w.walentukiewicz@ue.katowice.pl. ORCID: 0000-0002-2007-060X.

KEYWORDS: theory of basic concepts, overall similarity, deictic-inductive definition, categorization.

1. TERMINOLOGICAL REMARKS

This section will specify how certain key terms are to be understood. As a point of departure, let us note that categories are composed of objects from the world. The word "category" is equivalent to the terms "naïve set" (Voitsekhovskii, 1995, p. 120) and "range of a name". The expression "naïve set" (hereafter, simply "set") is defined on the basis of the similarity relation, and is not a term from set theory. However, categorization is a specific form of intellectual activity consisting in, amongst other things, an acknowledgement, on the basis of either overall perceptual similarity or some particular perceptual properties, that certain objects belong to one and the same given (naïve) set.

Objects can be concrete or abstract (the former are derived from perception, whilst the latter are not). Concrete objects are divided into concrete things (e.g. a dog), concrete properties (e.g. green) and concrete relations (e.g. taller). Abstract objects are divided into abstract things (e.g. dogness), abstract properties (e.g. greenness) and abstract relations (e.g. tallness). This article will not be dealing with abstract objects.

Objects s h o w n during categorization will be defined as "samples", while those subsumed u n d e r a n a m e will be called "semantic types". Categories are themselves formed around either samples or semantic types. In language names represent categories. The present analysis will focus on the names of *concreta*, not *abstracta*.

Names are also divided into general names (involving more than one, e.g. "dog"), singular names (involving just one, e.g. "Ajdukiewicz") and empty names (with none, e.g. "square wheel"). This article will not attempt to deal with singular or empty names.

A particular delimitation of a set of objects (e.g. those designated by the word "dog") is introduced by identifying properties possessed by those objects, e.g. barking, or by determining similarities to a selected object-sample (e.g. similarity to a dog that is pointed to as representative). Meanwhile, a delimitation of a set of properties (e.g. designated as "red") or relations (e.g. designated as "being taller") is introduced by identifying a similarity to a given sample of such a property or relation. For instance, a delimitation of the range of the set of red objects is introduced by identifying their similarity to a given sample of red, while the delimitation of the range of the set of relations of being taller is introduced by identifying their similarity to a sample of being taller (such as the relation that is John's being taller than Johnny, directly shown).

Semantic relations occurring between the realm of mind or language and elements of the world will be defined by the use of the term "reference" (Carey, 2009, p. 487). A reference contains naming, designating and standing for.

A similarity is distinguishable by the following: either it is reducible to definite common parts (generally referred to as properties) or it is not so reducible. In the former case, it will be possible to talk interchangeably about possessing a similarity (further referred to as "a detailed similarity") and possessing definite common properties—even though this risks being accused of falling foul of Ockham's razor by permitting the number of posited entities to be needlessly multiplied. In the latter case, it will not be necessary to discuss the existence of common properties, as this similarity (further referred to as "the overall similarity") will not be reducible to any definite properties.

We are of the opinion that some of the most important tools enabling the binding of elements of language or thought with elements of the world are deictic (ostensive) definitions. Ostensive definitions are elementary sentences of the form "This is *N*", in which the word "*N*" can be replaced by proper nouns, definite descriptions or general names. Such definitions bind names (or concepts) with individual objects or sets of objects.

Baptismal acts of references of first names, acts of presenting references of specific descriptions and the first acts of naming selected references of general names are made by means of deictic definitions. Such elementary sentences will hereafter be referred to as "simple deictic definitions". The inductive version of ostensive definition is built from multiple "This↓ is *N*" expressions (the initial condition), combined with a rule which allows for the determination of non-demonstrative objects (the inductive condition). This will be further referred to as "deictic-inductive definition" ("ostensive-inductive definition"), and can take two forms:

$$\text{this}\downarrow \text{ is } N \wedge \ldots \wedge \forall x(W(x) \rightarrow x \text{ is } N)$$

or

$$\text{this}\downarrow \text{ is } N \wedge \ldots \wedge \forall x(x \text{ is } N \equiv W(x)).$$

These schematizations will be utilized to describe the structure of basic concepts. In any such schema, *N* represents the name being defined (the definiendum), the variable *x* corresponds to the relevant set of objects, and *W* states the conditions which objects under consideration must fulfil if they are to count as falling within the range of the name so defined.

This leaves the issue of how concepts are to be defined. Here, we shall only appeal to an understanding of concepts of the most general kind. For Rosch (1978), concepts are mental representations of objects that are in some sense self-equivalent, while according to Murphy (2004) they are just mental representations of categories.

Contemporary psychological and philosophical literature mentions three paradigms of concepts: exemplars (where categories are formed around individual properties, i.e. such properties as determine a single designatum), prototypes (where categories are formed around typical properties), and what theory-based models propose (where categories are formed around theories explaining correlations between properties; see, e.g., Machery, 2009; Walentukiewicz, 2011). None of these notions of concepts takes into account overall similarity—which is irreducible to specific common properties—as an instrument of categorization.

The following concepts (Walentukiewicz, 2011) can thus be distinguished (as meeting the criterion of being constructed around relations or properties):

1. basic (formed on the basis of an overall similarity relation, as well as indistinguishability);
2. non-basic (formed on the basis of properties).

Basic and non-basic concepts can be bound to either general or singular names. This article will deal solely with the former.

2. BASIC-LEVEL CONCEPTS

In general, it is fair to say that the psychological literature does not aim to explore basic concepts as such, or enter into discussions regarding what the very notion of a basic concept amounts to. A theory of the latter kind has only been introduced in the context of philosophical research (Walentukiewicz, 2011), whereas what we find presented in the psychological literature is the notion of a basic-level concept (cf. Rosch, 1978). Therefore, before dealing with the former, we should first introduce the latter, with the aim of determining how it stands relative to the notion of a basic concept itself.

Rosch (1978) writes about concepts in term of their possessing horizontal and vertical dimensions. The former refers to ways in which concepts differ as to type (i.e. in respect of their building), whilst the latter deals with how they vary with regard to their degree of abstraction. Horizontally, concepts divide up in terms of building, whereas vertically they are distinguished by their range.

Our own approach here will be, on the one hand, to adopt Rosch's nomenclature when discussing the vertical dimension of concepts, and as a basis for introducing the term "basic-level concept" itself (implying an intention to discuss concepts in their vertical dimension), while on the other, to introduce the structure of a new type of concept (and hence open up possibilities for discussing concepts in their horizontal dimension). Basic-level concepts, after all, are themselves a type of concept, so they, too, must possess some sort of building.

We, therefore, advance the thesis that such concepts have a special building, which has not been addressed via any of the notions of concepts so far proposed in the literature, and the aim of the present article is to substantiate precisely this thesis. Given that the literature has adopted the adjective "basic" to characterize

a certain level of concepts, and concepts of this level are also taken to furnish a distinct type of concept, it seems reasonable to retain the same descriptor for the latter. That, in brief, is why we shall define "basic concept" in terms of the idea that what it actually refers to are basic-level concepts.

3. SELECTED PROPERTIES OF BASIC CONCEPTS

So far, we have only sought to introduce the terminology adopted from existing sources. Now, with a view to identifying some initial properties of basic-level concepts, we shall set out some observations. Indeed, our exclusive focus in the ensuing investigation will be on arriving at an adequate description of four specific properties pertinent to basic concepts.

3.1. Basic Concepts Are Formed by People Belonging to Various Cultures

Both anthropologists and psychologists distinguish several levels of categorization. These levels are differentiated according to the degree of abstraction involved. Concepts of the highest level possess the highest degree of abstractness, whereas concepts on the lowest level have the lowest. On the basis of research into ethnic cultures (e.g. the Tzeltal Maya tribe from Southern Mexico, or the Aguaruna Jivaro tribe from Northern and Central Peru), the anthropologist Brent Berlin (1978) enumerated the following levels of biological categorization: the kingdom (e.g. plant, animal), the living form (e.g. tree, fish), the indirect level (e.g. evergreen, fresh water fish), the generic level (e.g. pine, bass), the specific level (e.g. white pine, black bass) and the varietal level (e.g. Western white pine, large-mouthed black bass). Berlin established these levels through observational methods, noting what objects and sets are distinguished by Indian tribes from Central and South America, what names have been assigned to these objects and sets, and what subset-inclusion relations Indian tribes identify.

Meanwhile, the psychologist Eleanor Rosch (1978) differentiated three main levels in so-called Western culture: superordinate (e.g. tree), basic (e.g. oak) and subordinate (e.g. white oak). The experiment carried out by Rosch and her research team adhered to the following general outline: the test group consisted of subjects for whom English was their native language. Nine taxonomies were selected: of trees, birds, fish, fruit, musical instruments, tools, clothing, furniture and vehicles (these made up the superordinate level). Then, with reference to anthropological investigation (when dealing with biological categories) and the cognitive intuition of the authors of the experiment (when dealing with non-biological categories), the level of basic categorizations was established and designated as the basic level (e.g. sets of oaks, maples and birches, and sets of chairs, tables and lamps, made up the basic level). The sets belonging to the subordinate level (e.g. the set of standing lamps and the set of desk lamps, the set of white oaks and the set of red oaks) were contained within the set belonging to the basic level. Subjects had to know the words which these sets referred to. The

experiment participants were then asked to supply properties which they connected with categories of individual levels. The results were as follows: few common properties were listed for sets of the superordinate level (e.g. in regard to common properties of furniture or common properties of trees), whereas a multitude of differentiating properties were identified (e.g. properties which differentiate furniture from trees). With regards to basic-level sets, a comparatively large number of common properties were listed (e.g. for sets of chairs, tables and lamps), as were a similarly large number of differentiating properties (e.g. properties differentiating chairs, tables and lamps from each other). However, there was no fundamental increase in common properties listed amongst sets of the subordinate level (e.g. for sets of standing lamps and desk lamps) in comparison to that of the basic level, and relatively fewer differentiating properties were identified for sets of this level. On this basis, it was concluded that the level of common categories (the basic level) is particularly emphasized during the categorization process. The experiment also proves that the basic level is strongly embedded within the human categorization system. Human beings utilize it even after having mastered their language.

The basic level appears in the categorization processes of all cultures, or at least the majority of them (Rosch, 1978), and corresponds to Berlin's typological levels. Usually, the basic level discussed pertains to some object. Even so, in the available literature from the field of psychology, it is sometimes suggested that the basic level appears not only during the categorization of the object, but also during the categorization of its properties and relations. For example, the following levels of concepts can be distinguished for properties and relations (cf. Walentukiewicz, 2011):

- superordinate level: *colour*,[1] *relation*;
- basic level: *green*, *red*, *taller*, *shorter*;
- subordinate level: *Caucasian red*, *the redness of the lips*, *taller by about 5 cm*, *taller by about a head*.

George Lakoff (1987, pp. 270−271, 300), meanwhile, supplies additional examples of this sort for specific concept levels:

- superordinate level: *moving*, *ingesting*;
- basic level: *running*, *walking*, *eating*, *drinking*;
- subordinate level: *ambling*, *slurping*.

[1] The symbols *…* are used to indicate concepts.

Level-specific concepts of these kinds, on his account, emerge in connection with such properties of the basic level as *tall*, *short*, *hard*, *soft*, *heavy*, *light*, *hot*, and *cold*, as well as such relational concepts as *on*.

Although some of the available literature (cf. Murphy, 2004, p. 229) argues that the basic level for properties and for relations has not yet been adequately researched, for the purposes of this article, our position—without going into unnecessary details—will be that this concept level can be distinguished for some properties and some relations. Whereas there is little dispute regarding which properties or relations belong to the basic level, there are substantial differences when identifying examples belonging to the superordinate and subordinate levels.

3.2. Basic Concepts Undergo Initial Formation During Early Childhood

It has been experimentally proven that infants aged three or four months are able to form concepts for certain objects, properties and relations.[2]

Peter D. Eimas and Paul C. Quinn (Eimas & Quinn, 1994; Quinn & Eimas, 1996, p. 195) have established that three- or four-month-old infants are able to form basic-level concepts for such animals as cats or dogs. Children of this age group are also able to form sets for horses that exclude, for instance, cats, giraffes and zebras (dogs, cats, horses, giraffes and zebras belong to sets located on the basic level in adult classifications; cf. Behl-Chadha, 1996, p. 107).

Gundeep Behl-Chadha (1996) has also determined that three- or four-month-old children are able to acquire basic-level concepts for such objects as couches, chairs or beds. They proved capable of forming a set of couches that excluded chairs and beds (Behl-Chadha, 1996, p. 120), and a set of chairs that excluded couches and beds (chairs, couches, beds belong to sets located on the basic level in adult classifications (Behl-Chadha, 1996, p. 112, 115). The children established sets of chairs comprised of kitchen chairs, swivelling desk chairs, rocking chairs, and upholstered chairs, all varying in colour and style (Victorian, rococo, colonial, contemporary, etc.; Behl-Chadha, 1996, p. 113, 115−116).

The available literature (Bornstein, Kessen, & Weiskopf, 1976) supplies evidence to suggest that colours, despite being continuous in nature, are distributed by children between qualitatively deviating categories, at least in regard to four of them, these being red, green, yellow and blue (where these belong to a set of 11 colours designated by Berlin and Kaya [1968] as “basic colours”). Experiments have proven, further, that children distinguish the colours red, blue, green and yellow from one another (Bornstein et al., 1976, p. 201).

Monochromatic light does not exist under natural circumstances. One can only contemplate whether similar results would have been reached if children had been shown colours existing in the natural world. Light waves are rays, they do

[2] These concepts need not contain all of the potentially relevant designata.

not reflect light; however, the colours possessed by objects come into being as a result of the reflecting of light from their surface.

Anna Franklin and Ian R. Davies (2004) established that three- or four-month-old children not only distinguish basic colours (red, blue, green and yellow) formed as a result of the reflecting of light, but also distinguish primary colours from secondary ones[3] (in this case, blue from purple, or red from pink; Franklin & Davies, 2004, p. 375). Therefore, there is no significant difference between the results of their investigations and the results obtained by Bornstein et al. (1976).

The results of these experiments also furnish evidence to the effect that basic-level objects are the earliest observable type of objects.

Subsequent research, undertaken mainly by psychologists and linguists, has introduced slight alterations to the above findings. These assert that although there is generally no variation between different languages (cultures) regarding the most expressive colour examples, borderline examples do cause disputes. Language can blur the lines between basic colours through the use of the same or a different name for certain given samples: this causes differentiable colours to be included in one set, or a one-colour set to be broken up into two different sets (Franklin & Davies, 2004, p. 373; Wierzbicka, 1999, p. 405−449).

The existence of basic level concepts is evidenced by the experiments conducted by Paul C. Quinn (1994). In his opinion, three-month-old children are able to form concepts for relations such as *above* or *below*, because they are able to distinguish between a situation where a dot is placed above a rod and one where it is located beneath it (Quinn, 1994, p. 58−60): "[…] young infants can […] form categorical representations of physical space that are defined by the positional relations of objects in the environment" (p. 66).

Four-month-old children are not able to master a language. That is why these experiments also ascertain that concepts for sets of objects, properties and relations, at least in some cases, are formed before children learn a language.

3.3. Basic Objects Are the Earliest to Be Named

Rosch claims that "The basic level of abstraction is that level of abstraction that is appropriate for using, thinking about, or naming an object in most situations in which the object occurs […]" (1978, p. 43).

Many researchers (Bloom, 1993; Clark, 2003; Fenson et al., 1994) corroborate the claim that names for basic objects are the first to be learnt. Children between the sixteenth and twenty-fourth month use nouns such as "ball", "milk", "bread", "apple", "dog", "cat", "mommy", "daddy", and "nana", verbs such as

[3] According to Franklin and Davis, the division introduced by Berlin and Kay treated black, white, red, green, yellow and blue as primary colours, while brown, purplish red, pink, orange and grey were classed as secondary ones (Franklin & Davies, 2004, p. 350).

"cry", "come", "sleep", and "want", and names of properties such as "red" and "green". These words are single signs (Rosch, 1978, p. 35) that are "easy to say and remember" (Stern, 1959, p. 8; Berlin, 1994, p. 92).

The basic level is used by human beings even after they have mastered a language. This does not mean that human beings form basic concepts once and for all, or that they do not undergo modifications. On the contrary, they are subject to occasional changes.

3.4. Basic Concepts Are Inconstant

The basic level is not rigidly fixed: it is malleable. Lee Brooks (1978) draws attention to this flexibility: it can become attenuated as a result of a lack of interest in distinguishing certain sets, or enhanced through specialist training. In general, city-dwellers consider *tree* to be a basic category, and not *maple* or *larch*.[4] Horse or dog experts regard subordinate categories (e.g. breeds of horses or dogs) as basic. Rosch (1977, pp. 42−43) notes that for the majority of people, airplanes are treated as being basic-level, but the same does not apply to aircraft mechanics:

> […] he considered aeroplanes as a whole more similar to each other than vehicles […], aeroplanes appear to be an example of a category in which either one or two sets of correlational structures are available depending upon the degree of knowledge of the perceiver. (Rosch 1977, p. 43)

In other words, it is possible to "heighten" the basic level so that it becomes the superordinate level without expanding one's cognitive abilities, or conversely to "lower" the basic level, thus changing it into the subordinate level through training. Nevertheless, basic-level concepts, in the majority of cases, will be used by an individual human being throughout their entire life.

So what, then, is our current state of knowledge regarding basic-level concepts? The research of anthropologists and psychologists has delivered the following characterization of them (only properties relevant to our considerations are listed here):

a) They are present in the systems of classification of multiple diverse cultures.

b) Objects from this level are the first to be recognized (and observed): e.g. the initial observation is of a dog, not a mammal or a dachshund. Moreover, they are assimilated at an early phase.

---

[4] It should be noted that the term "tree" is not a term of botanical categorization, whereas the names "maple" and "larch" refer to genuses (Szweykowska & Szweykowski, 1993, p. 132; Browicz, 1993b; 1993c, p. 265, 386).

c) Objects from this level are the first to be named, most often using short words. During the initial phase of learning a language, parents and children use words relating to basic-level objects. These words are also utilized in neutral contexts, e.g. by saying "Look, a dog!" and not "Look, a mammal!" or "Look, a dachshund!" (as in the last two cases, specialist knowledge regarding mammals and dachshunds is required).

d) They are malleable: i.e. during the initial phase of learning a language, dogs are basic objects, whereas with time and after expanding one's knowledge about them, the dachshund subspecies may itself become a basic object. (People living in cities use the term "tree" on the basic level, instead of, say, the term "beech").

Even so, the above-mentioned properties do not in themselves seem informative about the building of basic-level concepts. At the same time, various authors have pointed to the need to describe sets belonging to the basic level (Lakoff, 1987; Arterberry & Bornstein, 2001), so before proceeding to our own investigations it makes sense to consider what the available literature has to say regarding the building of these concepts.

4. CRITICISM OF EXISTING HYPOTHESES

On what basis do children form categories of dogs, cats or horses? What are the predominant hypotheses regarding this in the available literature? An attempt to reach conclusions on the basis of the material available points to a substantive divergence in opinion amongst psychologists. The psychological literature suggests various grounds for differentiating basic-level categories: overall shape (e.g. Marr, 1982, pp. 215–233, 295–328; Landau, 2004, pp. 118–119), external head contours[5] (Quinn & Eimas, 1996, p. 189, 206–207), or specific properties such as a configuration of facial properties (Quinn & Eimas, 1996, p. 191, 209). Some authors specify elements which are not taken into account when differentiating categories belonging to the basic level: e.g. children do not utilize information regarding an animal's torso (Quinn & Eimas, 1996, p. 196). Several critical remarks can be formulated with regard to the above hypotheses.

Where shape is concerned, we should note that what a child actually observes are objects: e.g. dogs assuming various positions—standing, say, or lying down. The observed shape differs from position to position. Does a child remember each of these shapes? A positive answer to that question will surely elicit the observation that there can be many such positions, and that the memorizing of each and every one of them by a three- or four-month-old infant would excessively burden the child's memory. Furthermore, a child would have to remember

[5] Hereafter, the shape as well as the head contours will be recognized as specific object features.

many different additional animal shapes (e.g. those exhibited by a cat), and then be able to distinguish each shape of one animal from every shape of another animal—for instance, the various positions of a cat and a dog, respectively. At the same time, we should note that any issues surrounding the role played by shapes will be inapplicable when it comes to explaining the categorization of properties and relations.

Quinn and Eimas also put forward the claim that animals such as dogs and cats are recognized and distinguished by their external facial properties (Quinn & Eimas, 1996, p. 200). However, a dog's face is not the dog itself. In order to define the dog, the whole animal should be presented. Supplying a dog's face for observation would most likely result in the formation of the concept *dog's face*. If such an object existed independently of the dog, then a dog's face would have to, say, fly around like a wasp. Are children, therefore, really unable to distinguish a dog from its face? To conclude, we may say that it is hardly feasible to reduce the defining of dogs *per se* to a mere presenting of parts of a dog, as in the example of the dog's face.

The above-mentioned authors also state that they are unable to supply properties—including those other than the aforementioned—to which children could refer in order to distinguish cats, dogs and female lions (Quinn & Eimas, 1996, p. 191). The authors ascertain the following:

> With respect to the internal features, we do not know whether the categories are specified by the dimensions of one particular feature (e.g., the nose) or by the configuration of several features (eyes, nose, mouth). Similarly, we are unable to say whether it is the "gestalt" of the outer contour or some specific region of the external border of the head […]. (Quinn & Eimas, 1996, p. 208)

Hence, what we find is that the existing literature does not conclusively explain the issue of the formation of these categories, in that doubts arise as to the hypotheses formulated, due to the actual state of our knowledge pertaining to this topic. It remains simply unclear what type of information infants are able to use during categorization (Arterberry & Bornstein, 2001, p. 334).

Nevertheless, such speculative theorizing is not without a certain significance: firstly, it draws attention to the problem of the impossibility of delimiting the range of some category without any prior delimitation of the relevant properties, while secondly, it raises the issue of the impossibility of delimiting the properties qualifying for membership of that category without any prior delimitation of its range. Thus far, the available literature mentions two methods for distinguishing categories: on the basis of similarity and on the basis of properties (Walentukiewicz, 2011). Let us, then, now introduce the potentially controversial supposition that infants aged three- or four-months do not perceive similarities. Therefore, we suppose, the child can only use properties for categorization. Yet in order to establish some properties, the infant must presumably have prior knowledge of at least some category exemplars, etc. (given the problem highlighted above, regarding the impossibility of delimiting the range of a category without a prior

delimitation of the relevant properties, and the impossibility of delimiting properties which qualify for membership of the category without prior delimitation of its range). Even if the infant establishes a preliminary range, how is he or she supposed to know which properties determine inclusion in the category, and which exclude it from the latter? Comparing, in this regard, exemplars that belong to the category with others that do not would, in practice, take so long that the child would become discouraged from exploring the subject matter (since an infinite amount of time would be needed to compare the objects in order to successfully delimit the properties). Without the possibility of using overall similarity, three- or four-month-old children would be unable to form concepts for dogs or wardrobes.

If these concepts were formed on the basis of properties, then they would not be malleable: if the object were to possess definite properties, and these properties decided about its belonging to a category of a given level, then our perception of this fact would not be alterable. Let us assume, say, that possession of the plant-specific fruit called "acorns" is the distinctive property of oaks. Under our current supposition, this property will determine its belonging to a given genus. Nonetheless, it does not determine whether the exemplar in question belongs to any particular subspecies, such as the cork-oak (where this would require introducing further properties, e.g. the possession of a thick and elastic bark, and so on), or family (since possessing acorns is not a distinguishing property for the beech family; see Browicz, 1993a, pp. 118−119). Equally, what determines that some animal is a giraffe is its individual DNA code. This property distinguishes the species, but not subspecies of giraffe (giraffes have the same DNA, but differ in respect to pigmentation details and horn quantity) or the ruminant family. With respect to overall similarity, on the other hand, the situation differs entirely: initially, elephants can be assigned to one basic-level set, but perceptual experience will subsequently allow for this to be divided into the subsets made up of African and Indian elephants, respectively. The similarity internal to the African elephant subset, along with that internal to the Indian elephant subset and taken together with the dissimilarity between African and Indian elephants, will all be sufficiently significant in their own terms to mean that perception alone allows for the two categories to be distinguished, without any need to establish their distinctive properties.

It is generally assumed that children, at the moment of birth, do not possess any knowledge of the outside world. Essentially, they must acquire any such knowledge.[6] This is why conceptual content has to be arrived at on the basis of perception, at least during the initial formative phase of concept creation. If the child does not possess knowledge about the world, especially in regard to the distinctive properties of a category, then it must be the case that he or she utilizes

[6] This by no means implies that a child enters our world as a *tabula rasa*. He or she possesses principles which determine categorizations (principles specified by Hirsh-Pasek, Golinkoff, Hennon, & Maguire, 2004). These principles will not be discussed in this paper.

some other tool, which must be accessible via perception, available to children aged three- to four-months or more, and such as not to require significant cognitive abilities. Three- or four-month-old children are incapable of performing subtle analyses in order to establish common properties. Hence they form primary categories "without close study" (Berlin, 1978, p. 15).

The literature to date discusses both similarity and properties as possible means for identifying categories. (We should add that the theories put forward generally seek to determine which properties are important for a given field, whereas what concerns us here is not the validity of such theoretical claims, but rather what the real status is of any properties identified as such. (Walentukiewicz, 2011). Since three- or four-month-old children are unable to use properties during categorization, overall similarity is the only tool left at their disposal. By excluding properties, similarity remains as the only possible means of categorization.

5. OVERALL SIMILARITY

To begin with, let us sketch an example of a relation of overall similarity. When observing a herd of elephants and a herd of giraffes traversing, is it really necessary to refer to properties in order to determine at first glance that elephants belong to one category, and giraffes to another? What allows human beings to appropriate elephants to one category and giraffes to a different one? Given that a child does not recognize properties that could determine that something belongs to either the *elephant* or the *giraffe* category, similarity must be the sole remaining option. So what principles govern this process?

Let us first of all distinguish categorization from ordering—and with this, similarity, which is a categorizing relation, from whatever would count as an ordering relation.[7] By asserting that a given exemplar (e.g. a sparrow) constitutes a better example of a given concept (e.g. *bird*), since it shows closer similarity to a selected sample (e.g. a thrush) than to a selected alternative (e.g. a penguin), an ordering is introduced. (Hence, the similarity is a three-argument relation: *x* is more similar to *y* than to *z*.) Conversely, by asserting that the exemplar (e.g. the sparrow) belongs to the concept-range at issue (e.g. that of *bird*) by virtue of its similarity to an identified sample (e.g. thrush), a categorization is introduced. (Here the similarity is a two-argument relation: *x* is similar to *y*.)

Above and beyond this, categorizations of concepts can be subdivided into those formed by appeal to detailed similarity, and those arrived at on the basis of overall similarity.

[7] Categorization should also be distinguished from classification. Classification establishes relationships between categories and is in conformity with logical modes of division − as fulfilling the conditions of being both exhaustive and separable. Categorizations can, of course, lead to classifications.

Detailed similarity (i.e. similarity in some respects and to some extent[8] (Kotarbińska, 1959, p. 47), reduced to the possession of common properties, can be presented using the following schema of Ajdukiewicz (1975):

$$\forall_{x,y}[xP_Cy \equiv C(x) \wedge C(y)],$$

in which x represents a certain object, and P_C represents the relation of similarity reduced to possession of a common property C. The schema can be interpreted as follows: for every x and y, x will be similar to y with respect to possession of the property C if and only if x and y possess the property C. (Thus, for example, a polar bear and a refrigerator will be similar to one another with respect to the colour white, if and only if the polar bear and the refrigerator are both white.) According to this schema, the relation of detailed similarity is an equivalence relation. This, in turn, will be a reflexive relation:

$$\forall_x (xP_Cx).$$

This schema is to be interpreted as follows: for every x, x will be similar to x with respect to possession of property C (e.g. a polar bear will be similar to itself with respect to its possession of white fur).

This will also be a symmetric relation:

$$\forall_{x,y}(xP_Cy \rightarrow yP_Cx).$$

Here the schema is interpreted thus: for every x and y, if x is similar to y with respect to the possession of property C, then y will be similar to x with respect to possession of that same property (e.g. if a polar bear is similar to a refrigerator with respect to being white, then the refrigerator will be similar to the polar bear with respect to this colour).

Furthermore, this is also a transitive relation:

$$\forall_{x,y,z}(xP_Cy \wedge yP_Cz \rightarrow xP_Cz).$$

[8] On the one hand, Kotarbińska (1959) recognizes similarity as being in some respects and to a certain extent an intransitive relation (p. 65), while on the other hand she claims that such similarity is reducible to definite common properties (p. 47). If an exemplar possesses common properties, then all exemplars in the category must be similar to one another with respect to these properties and such a similarity must be a transitive relation. Yet this raises the question of whether it is conceivable that an object could possess a feature that would conclusively assign it to a given category, even when this object was in no way similar to a certain exemplar of this category that also possessed this same feature, and where the similarity consisted in a relation derived from possession of the latter. Such a scenario seems entirely implausible.

In the above instance, we interpreted the schema in this way: for every *x*, *y* and *z* , if *x* is similar to *y* with respect to possession of property *C* and *y* is similar to *z* with respect to possession of property *C*, then *x* will be similar to *z* with respect to possession of property *C* (e.g. if a polar bear is similar to a refrigerator with respect to being white, and a refrigerator is similar to a washing machine with respect to being white, then the polar bear will be similar to the washing machine with respect to being white).

Now we can ask, does overall similarity possess such logical characteristics? Giraffes, for instance, at least with regard to adults and non-defective specimens, are generally similar to one another. We are able to combine them into one set without specifying definite properties.

Overall similarity in respect of the category *giraffe* is a reflexive relation: each giraffe is generally similar to itself. It is also a symmetric relation: if one giraffe is generally similar to another, then the second will be generally similar to the first. Such similarity is not an asymmetric relation (as might be inferred from the remarks concerning similarity made by Tversky [1977]). Asymmetry is rather a feature of the ordering relation, which can be defined as follows:

$$\forall_{x,y}(xPy \rightarrow \sim(yPx)).$$

When forming its first basic-level categories, a child would have to be aware of the ordering criterion. Moreover, the ordering of the category would require the child to have pre-constructed the given category, or at least part of it. Yet category formation precedes ordering. Therefore, the similarity mentioned by Tversky (1977) cannot refer to the formation of primary categories.

It would appear that overall similarity is also an intransitive relation: even if one object is similar to another (e.g. a son is similar to his father), and this second is similar to a third (e.g. the father is similar to his father), the first need not be similar to the third (e.g. the son need not be similar to his grandfather). So is overall similarity really therefore intransitive?

Let us analyse the tools used by three- or four-month-old children during categorization. One argument could be as follows: all exemplars falling within the *giraffe* category are similar to each other. Overall similarity is a transitive relation: there are no two exemplars of *giraffe* which would not be similar overall to each other. Meanwhile, a different argument could run as follows: within the basic category *dog*, dachshunds are similar to one another, but they are not similar to a husky. Within the category *dog*, it is possible to find such exemplars as will not be similar to one another: in other words, starting out from a representative of one breed, and proceeding in terms of overall similarity, it is conceivable that a representative of another breed could be found that would be dissimilar to the first one. Hence, it can be concluded that overall similarity must be an intransitive relation. Yet the same relation cannot possess both of these mutually exclusive features. So there is a problem here that we need to at least try to elucidate.

We should note that the logical character of a relation is something altogether different from its functional operation within a category. More specifically, the relation of overall similarity is an intransitive one, yet it nevertheless functions within certain categories as a transitive relation. Natural discontinuities which exist in the biological world between certain basic-level categories are quite efficacious in this regard (this remark refers to basic categories and not species as such[9]—species themselves only occasionally constitute basic-level categories) and, for example, protect animals from different categories from mating. Anthropologists (e.g. Berlin, 1978), biologists (e.g. Mayr, 1984, pp. 531–540) as well as palaeontologists (e.g. Eldredge, 1995) discuss the subject of natural discontinuities. Within these categories, overall similarity functions as a transitive relation.

A three- or four-month-old child is observing giraffes. He or she does not know any names, but is required (or just wants) to form his or her first category. Giraffes are practically indistinguishable from one another. Each giraffe (when full-grown and without defects) is similar overall to other giraffes. If several distinct objects, in this case giraffes, exhibit overall similarity to one another, then this will suggest to the child the existence of a category to which they all belong, while, conversely, if they do not show any such overall similarity to one another, this will suffice to mean that no such category exists. The above reasoning can be formulated be means of the following schema:

$\forall z_1 \forall z_2[(z_1 \neq z_2 \wedge z_1 P z_2) \rightarrow \exists Z(z_1 \in Z \wedge z_2 \in Z)]$ (Koj, 2007, p. 65),

$\forall z_1 \forall z_3[(z_1 \neq z_3 \wedge \sim(z_3 P z_1)) \rightarrow \sim \exists Z(z_1 \in Z \wedge z_3 \in Z)]$.

Here, the variables z_1, z_2 and z_3 represent members of the relevant category of objects, the symbol P expresses the relation of overall similarity, and the variable Z refers to the category in question.

Within these categories, the relation of overall similarity will be one that closes the category. So the category Z is closed by D with respect to the relation S ($D(Z, S)$), when each exemplar standing in relation S to the sample belonging to category Z also itself belongs to category Z:

$D(Z, S) \equiv \forall z_1 \forall z_2[(z_1 \in Z \wedge z_2 S z_1) \rightarrow z_2 \in Z]$.

[9] The polar bear (*Ursus maritimus* or *Thalarctos maritimus*) is descended from the brown bear (*Ursus arctos*). It is widely accepted that they belong to two distinct species. Recently, polar bears have been seen to mate with grizzly brown bears (a sub-species of the brown bear—*Ursus arctos horribilis*). From such a relationship then issue so-called grolar bears, leading scientists to wonder whether these represent a new species. The polar bear differs from the brown bear only in terms of the colour of its fur. They are similar in weight and in body-length, as well as in their possessing a noted predisposition for swimming. This would suggest that natural discontinuities do not appear at the species level. Yet this particular case does not contradict the thesis that they exist as such: it is not possible to cross a bear with a salmon, as the bear would sooner eat the salmon. (Here we speak in jest!)

This, indeed, is how primary categories are formed. When a child begins to learn a word, he or she will already use such categories and attach names to them. A child is capable of forming a wide array of such categories, as there are many species, kinds or even races in the world, between which a sufficient number of natural discontinuities exist to allow for a fairly straightforward assignment of similar exemplars to one category, and differing exemplars to diverse categories, all on the basis of overall similarity.

When the first names begin to appear, the teacher can influence the learner to modify his or her primary categories by joining together objects that are dissimilar to one another into one category: e.g. the teacher can group dachshunds and huskies together into one category, naming them with the word "dog", simply by employing deictic definitions. Equally, the teacher can also exclude objects possessing overall similarity to one another (e.g. wolves and huskies) from co-membership of the same category, giving them different names, also by means of simple deictic definitions.

Only then can the intransitivity of the overall similarity relation reveal itself. Dachshunds are similar overall to one another, and huskies are similar overall to one another. Within the range of dachshunds or of huskies, this similarity works as a transitive relation. However, it functions as a non-transitive feature across the entire category of dogs. This can be formulated as follows:

P is an intransitive relation in category N (e.g. dogs) $\equiv \exists_{x, y, z \in N} (xPy \wedge yPz \wedge \sim(xPz))$.

When introducing its first names, the child need not rely on using properties in order to determine membership of an object in relation to a given category. Overall similarity, along with a capacity to include within one and the same category objects that are dissimilar overall on the basis of the fact that they possess common names, will prove sufficient—at least at this point in their development.

It can be assumed that overall similarity functions between wholes. The theses of Gestalt psychology furnish a theoretical background for this contention: e.g. the Wertheimer principle, which relates to groupings of objects of identical appearance or wholly similar objects within one category (Wertheimer, 1923; Palmer, 2002, pp. 101–102). Wholes are simpler than mere summations of parts (Lakoff, 1987). A single feature, or even several properties, will prove insufficient as a basis for a child's assigning objects possessing these properties to a single category. An overwhelming majority of common properties must be present. Each feature is relevant to determining such similarity—even such properties as they share with exemplars of different categories. Of course, the more common properties the objects possess, the greater will be the similarity (Tversky & Gati, 1982, p. 125). For this reason, we can partially agree with Murphy, who states that:

> […] the holistic similarity of old to new items is probably very important […], when members of a category are *all* holistically similar, there may be less reliance on memory of individual exemplars and greater reliance on a summary prototype. (Murphy, 2004, p. 85)

This same author stresses how relations of similarity hold between simplified objects, whereas, in our opinion, quite the converse is what is important: it is the similarity of whole objects that is the key factor.

The introduction of semantic types is of particular importance. These are objects which have been assigned a name, and with which other objects are compared, in order to establish whether the latter can or cannot be regarded as falling within the range of that name. They are introduced using simple deictic definitions. Where overall similarity is sufficient for category formation (as with, say, the relation of overall similarity forming the category of elephants), only one semantic type should be formulated. However, when overall similarity proves insufficient, other names allowing us to introduce multiple semantic types will be necessary. Within the category of dogs, the relation of overall similarity obtains between the semantic type and the exemplar under evaluation—e.g. between the dachshund type and other dachshunds—and in such a way that each exemplar from the breed is, in fact, suited to functioning as a semantic type. Thereafter, it is possible to assign the same name to multiple semantic types.

Basic categories are delimited in ways that are physiologico-linguistically determined. On the one hand, names are the first category designata to be assigned, while on the other, this allows for certain subcategories formed on the basis of overall similarity to be grouped together into one category, despite there being no overall similarity between them: e.g. dachshunds have an overall similarity to one another, huskies have an overall similarity to one another, but dachshunds are dissimilar to huskies; yet if we name both subcategories "dog", they end up falling into the same category. A teacher, by using an appropriate name, can unite certain subcategories formed on the basis of overall similarity into a single category, within which overall similarity does not in fact hold between all of its elements. Basic names are used in a neutral context.

The psychological and philosophical literature dealing with similarity discusses the notion of detailed similarity construed as reducible to common properties, regardless of whether these properties are theory-specific ones (Medin, 1988), typical ones (Rosch, 1978), logical ones (in the sense of being jointly necessary and sufficient for set membership) (Ajdukiewicz, 1975), or modally rigid ones (i.e. jointly necessary and sufficient for membership of some set in every possible world *W*) (Putnam, 1996; 1998, pp. 119−120; Walentukiewicz, 2011). This similarity is an equivalence relation. We, on the other hand, are proposing a different kind of similarity, irreducible to common properties and constitutive of a non-equivalence relation. Granting the correctness of such a thesis**,** contemporary theories of concepts—be they classical, prototype-based, theory-dependent or focused on modal rigidity—will be unable to explain the process of forming basic concepts. This leaves the theory of conceptual exemplars, on the

basis of which some authors (e.g. Machery, 2009) have claimed similarity to be reducible to individual properties. Yet it is the present author's view that such a theory refers only to singular concepts, and not to general ones. (Support for this thesis can be found in [Walentukiewicz, 2011].) Since we are exclusively concerned here with general concepts, and not singular ones, similarity of the kind that is reducible to individual properties will not be discussed.

The first basic categories, in the sense of those initially formed and then subsequently corrected through the introduction of appropriate names (the same names for objects belonging to the same category, and different ones for objects belonging to distinct categories), can be described with the use of the following deictic-inductive definitional schema:

$(a_1 \text{ is } N \wedge a_2 \text{ is } N \wedge \ldots \wedge a_n \text{ is } N \wedge b_1 \text{ is not } N \wedge b_2 \text{ is not } N \wedge \ldots \wedge b_m \text{ is not } N) \wedge \forall y \forall v \forall x[(y \text{ is } N \wedge xPy \wedge v \text{ is not } N \wedge \sim(xPv)) \rightarrow x \text{ is } N]$.

In the above, a_1, a_2 and a_n (individual constants) are standing for exemplars that are similar overall to one another and constitute good positive examples. Meanwhile, b_1, b_2, b_m are standing for good negative examples. Thus, we have, say, the colour of succulent grass, considered as a good example for the name "green", but we also have the cloudless sky, considered not just as a good example for the name "blue" but also as a good negative example for the name "green". At the same time, the variable *y* is an individual variable (it being possible to substitute an individual term (individual name) in place of such a variable) corresponding to a set of previously determined good samples, the variable *x* corresponds to the set of exemplars to be evaluated with regard to membership in the sense of falling or not falling within the range of the name defined, and the variable *v* corresponds to a set of good negative examples. It is necessary to remember that the basic level is delimited on the basis of at least two factors: overall similarity (the perceptual aspect) and the names employed (the linguistic aspect). It is therefore perceptual-linguistic in nature.

SUMMARY

Basic concepts are formed on the basis of overall similarity, which is a reflexive, symmetric and intransitive relation. In certain categories, however, this relation can function as transitive, due to the existence of natural discontinuities obtaining between categories of the basic level.

Some readers may be struck by the fact that we have not sought here to take into account the notion of family resemblance, as proposed by Wittgenstein (1997). Our reason for not doing so is that while the author of *Philosophical Investigations* wrote about similarity as an intransitive relation, he also took it to be reducible to an alternative—namely, bundles of features (Koj, 1969; 1988). Therefore, similarity, on Wittgenstein's account, cannot be overall similarity. Apart from this, we may note the following conclusions:

- human beings begin to apply overall similarity as a tool of categorization at the age of three or four months (and the assumption that they use it throughout their whole life seems justified);
- overall similarity is malleable;
- our first names for things are assigned to categories distinguished on the basis of overall similarity;
- overall similarity and overall dissimilarity are both perceptual: e.g. similarity is perceivable between giraffes, while dissimilarity is perceivable between giraffes and elephants;
- taking into account the claims of Gestalt psychology, overall similarity may be said to obtain between wholes; e.g. children recognize such a huge similarity between giraffes (it being unnecessary to identify common properties), and such a huge dissimilarity between giraffes and elephants (it being unnecessary to identify differentiating properties), that they are able to create categories of giraffes and elephants without any difficulty (there being no issue of which properties to select).

Since giraffes possess common perceptual properties with respect to almost all of their characteristics, and since natural discontinuities between categories of the basic level do not allow for exemplars belonging to separate categories to mate (procreate), a wide array of common properties is maintained within these categories. The categorizations formed on the basis of overall similarity will only ever be "breached" partially during the language-learning period, through the adoption of the same or a different name. When the teacher puts forward a semantic type, he or she names an exemplar—one which is not similar overall to the exemplars to which this name is assigned. Overall similarity within a subcategory delimited by semantic type remains a transitive relation, though it is an intransitive relation within the category as a whole. Overall similarity, and our first names for things, are the "tools" which allow for the formation of basic concepts.

Finally, we may also add that when children form basic-level categories, they are then able to compare exemplars in order to go on to establish properties (Gentner & Namy, 2004). Above all, on the account given here, basic concepts are no more and no less than mental representations of basic-level categories.

REFERENCES

Ajdukiewicz, K. (1975). *Logika pragmatyczna*. Warszawa: Państwowe Wydawnictwo Naukowe.

Arterberry, M. E., Bornstein, M. H. (2001). Three-Month-Old Infants' Categorization of Animals and Vehicles Based on Static and Dynamic Attributes. *Journal of Experimental Child Psychology*, *80*(4), 333−346.

Behl-Chadha, G. (1996). Basic-Level and Superordinate-Like Categorical Representations in Early Infancy. *Cognition*, *60*(2), 105−141.

Berlin, B. (1978). Ethnobiological Classification. In: E. Rosch, B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 9−26). Hillsdale, New Jersey: Lawrence Erlbaum.

Berlin, B. (1994). Evidence for Pervasive Synesthetic Symbolism in Ethnozoological Nomenclature. In: L. Hinton, J. Nichols, J. J. Ohala (Eds.), *Sound Symbolism* (pp. 76−93). Cambridge: Cambridge University Press.

Berlin, B., Kay, P. (1969). *Basic Colour Terms. Their Universality and Evolution.* Berkeley, CA: University of California Press.

Bloom, L. (1993). *The Transition from Infancy to Language. Acquiring the Power of Expression*. New York: Cambridge University Press.

Bornstein, M. H., Kessen, W., Weiskopf, S. (1976). Color Vision and Hue Categorization in Young Human Infants. *Journal Experimental: Human, Perception, Performance*, *2*(1), 115−129.

Bornstein, M. H., Kessen, W., Weiskopf, S. (1976). The Categories of Hue in Infancy. *Science*, *191*(4223), 201−202.

Brockman, J. (Ed.). (1995). *The Third Culture*. New York: Touchstone.

Brooks, L. (1978). Nonanalytic Concept Formation and Memory for Instances. In: E. Rosch, B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 169−211). Hillsdale, New Jersey: Lawrence Erlbaum.

Browicz, K. (1993a). Dąb. In: A. Szweykowska, J. Szweykowski (Eds.), *Słownik botaniczny* (pp. 118−119). Warszawa: Wiedza Powszechna.

Browicz, K. (1993b). Klon. In: A. Szweykowska, J. Szweykowski (Eds.), *Słownik botaniczny* (pp. 264−267). Warszawa: Wiedza Powszechna.

Browicz, K. (1993c). Modrzew. In: A. Szweykowska, J. Szweykowski (Eds.), *Słownik botaniczny* (p. 386). Warszawa: Wiedza Powszechna.

Carey, S. (2009). *The Origin of Concepts*. Oxford, New York: Oxford University Press.

Clark, E. V. (2003). *First Language Acquisition*. Cambridge: Cambridge University Press.

Eimas, P. D., Quinn, P. C. (1994). Studies on the Formation of Perceptually-Based Basic-Level Categories in Young Infants. *Child Development*, *65*(3), 903−917.

Eldredge, N. (1995). A Battle of Words. In: Brockman J. (Ed.), *The Third Culture* (pp. 119−128). New York: Touchstone.

Fenson, L., Dale, P. S., Reznick, J.S., Bates, E., Thal, D. J., Pethick, S. J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), 1–173.

Franklin, A., Davies, I. R. L. (2004). New Evidence for Infant Colour Categories. *British Journal of Developmental Psychology*, *22*(3), 349−377.

Gentner, D., Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In: D. G. Hall, S. R. Waxman (Eds.), *Weaving a Lexicon* (pp. 533−568). Cambridge, Massachusetts−London: The MIT Press.

Hall, D. G., Waxman, S. R. (Eds.). (2004). *Weaving a Lexicon*. Cambridge, Massachusetts−London: The MIT Press.

Hinton, L., Nichols, J., Ohala, J. J. (Eds.). (1994). *Sound Symbolism*. Cambridge: Cambridge University Press.

Hirsh-Pasek, K., Golinkoff, R. M., Hennon, E. A., Maguire, M. J. (2004). Hybrid Theories at the Frontier of Developmental Psychology: The Emergentist Coalition Model of Word Learning as a Case in Point. In: D. G. Hall, S. R. Waxman (Eds.), *Weaving a Lexicon* (pp. 173−204). Cambridge, Massachusetts−London: The MIT Press.

Koj, L. (1969). On Defining Meaning Families. *Studia Logica*, *25*, 78−87.

Koj, L. (2007). O początkach teorii zbiorów. *Studia Semiotyczne*, *26*, 59−80.

Koj, L. (1988). Schematy definicyjne dla rodzin znaczeniowych. In: Z. Muszyński (Ed.), *O nieostrości* (pp. 303─317). Lublin: Uniwersytet Marii Curie-Skłodowskiej.

Kotarbińska, J. (1959). Tak zwana definicja deiktyczna. In: Kotarbiński T., Kotarbińska J. (Eds.), *Fragmenty Filozoficzne. Seria druga*. *Księga pamiątkowa ku uczczeniu czterdziestolecia pracy nauczycielskiej w Uniwersytecie Warszawskim profesora Tadeusza Kotarbińskiego* (pp. 44−74). Warszawa: Państwowe Wydawnictwo Naukowe.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. Chicago−London: The University of Chicago Press.

Landau, B. (2004). Perceptual Units and Their Mapping with Language. How Children Can (Or Can't?) Use Perception to Learn Words. In: D. G. Hall, S. R. Waxman (Eds.), *Weaving a Lexicon* (pp. 111−148). Cambridge, Massachusetts−London: The MIT Press.

Machery, E. (2009). *Doing Without Concepts*. Oxford: Oxford University Press.

Marr, D. (1982). *Vision*. New York: W. H. Freeman and Company.

Murphy, G. L. (2004). *The Big Book of Concepts*. Cambridge, Massachusetts−London: The MIT Press.

Palmer, S. E. (2002). Perceptual Grouping: It's Later Than You Think. *Current Directions in Psychological Science*, *11*(3), 101−106.

Putnam, H. (1996). The Meaning of "Meaning". In: A. Pessin, S. Goldberg (Eds.), *The Twin Earth Chronicles. Twenty Years of Reflection on Hilary Putnam's "The Meaning of 'Meaning'"* (pp. 3−52). Armonk, New York−London: M. E. Sharpe.

Putnam, H. (1998). *Wiele twarzy realizmu i inne eseje*. Warszawa: Państwowe Wydawnictwo Naukowe.

Quinn, P. C. (1994). The Categorization of Above and Below Spatial Relations by Young Infants. *Child Development*, *65*(1), 58−69.

Quinn, P. C., Eimas P. D. (1996). Perceptual Cues That Permit Categorical Differentiation of Animal Species by Infants. *Journal of Experimental Child Psychology*, *63*(1), 189−211.

Rosch, E. (1977). Human Categorization. In: N. Warren (Ed.) *Advances in Cross-Cultural Psychology* (Vol. 1, pp. 1−49). London: Academic Press.

Rosch, E. (1978). Principles of Categorization. In: E. Rosch, B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27−48). Hillsdale, New Jersey: Lawrence Erlbaum.

Rosch, E., Lloyd, B. B. (Eds.). (1978). *Cognition and Categorization.* Hillsdale, New Jersey: Lawrence Erlbaum.

Stern, W. (1959). The Background of Linnaeus' Contributions to the Nomenclature and Methods of Systematic Biology. *Systematic Zoology*, *8*(1), 4−22.

Szweykowska, A., Szweykowski, J. (Eds.). (1993). *Słownik botaniczny*. Warszawa: Wiedza Powszechna.

Tversky, A. (1977). Features of Similarity. *Psychological Reviev*, *84*(4), 327−352.

Tversky, A., Gati, I. (1982). Similarity, Separability, and the Triangle Inequality. *Psychological Review*, *89*(2), 123−154.

Voitsekhovskii, M. I. (1995). Set. In: M. Hazewinkel (Ed.), *Encyclopaedia of Mathematics* (Vol. 5, pp. 120−121). Dordrecht−Boston−London: Kluwer Academic Publisher.

Walentukiewicz, W. (2011). *Definicje deiktyczne a pojęcia. Badania z pogranicza filozofii języka i psychologii kognitywnej*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Wertheimer, M. (1923). Unterschungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, *4*(1), 301−350.

Wierzbicka, A. (1999). *Język - umysł - kultura*. Warszawa: Państwowe Wydawnictwo Naukowe.

Wittgenstein, L. (1997). *Philosophical Investigations*. Oxford: Blackwell Publishers Ltd.