Natural Language and Logical Consequence

An Inferentialist Account

MSc Thesis *(Afstudeerscriptie)*

written by

Simon Dominik Vonlanthen
(born October 15th, 1993 in Frauenfeld, Switzerland)

under the supervision of **Dr Luca Incurvati** and **Dr Julian Schlöder**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
July 5, 2021	Prof Dr Yde Venema (chair)
	Dr Luca Incurvati (supervisor)
	Dr Julian Schlöder (supervisor)
	Prof Dr Arianna Betti
	Dr Paul Dekker
	Dr Karolina Krzyżanowska

Institute for Logic, Language and Computation

Abstract

This thesis investigates the relationship between natural language and formal logic. It is usually thought that natural language and formal logic come apart at some point, which might stand in tension with the reliance on natural language examples in logic research. Concerning the former point, Glanzberg (2015) recently argued that natural language neither generates a logical consequence relation, nor does it come pre-equipped with a distinction between logical and non-logical vocabulary. I start by expounding Glanzberg's arguments in detail, before defending them against the received criticism so far. On these grounds, I suggest that the issue lies with the referentialist assumptions of Glanzberg, and thus the idea of use-theoretic approaches to linguistic meaning is introduced. Specifically, I survey the idea of inferentialism, both concerning semantics as well as metasemantics. On the basis of these meaning-theoretic assumptions, I further motivate and excavate corresponding proof-theoretic conceptions of logical notions. At last, I compare whether the inferentialist meaning-theoretic assumptions square badly with proof-theoretic assumptions, in the same vein as Glanzberg's referentialist truth-conditional stance mismatched with the corresponding model-theoretic conceptions. I argue that there is no such mismatch for the inferentialist on the conceptual level, and that there are, moreover, reasons to be optimistic that we can, in fact, extract a logical consequence relation in natural language via pre-existing logical constants.

Acknowledgements

Writing a Master's thesis is no small task, and it cannot be properly done without the help of others. As such, I would like to take this opportunity to thank everybody who helped me throughout this time, be it directly involving the thesis itself, or by simply giving me a roof over my head for the duration. I would like to give my special thanks to both of my supervisors, Dr. Luca Incurvati and Dr. Julian Schlöder, who accompanied me throughout the entire creation process. Without their continuous suggestions, criticisms, and pointers to the literature – as well as the occasional pep talk and encouragement – the thesis would have turned out quite differently. I have thoroughly enjoyed working together, and I hope the feeling was mutual. Furthermore, my thanks go to the defense committee for the organisation of my public defense and the stimulating discussion during the session. I would also like to thank some of my fellow students, especially Bas Kortenbach, Lukas Zenger, Tianwei Zhang and everyone else who had to listen to me ramble on about philosophy, ordinary language and logic one too many times. Special thanks on the side of emotional support go to my friends and family for the much needed diversions, and last but not least to Bingzi Jiang for the aforementioned roof as well as the general patience for putting up with me during these stressful times. 我爱你!

Contents

1 Introduction

The literature on the nature of logical consequence as well as its various epistemological and metaphysical underpinnings is vast. It ranges from issues stemming from the relation between material and formal consequences, to the debate between model-theoretic and proof-theoretic conceptions,[1] all the way to the number of consequence relations out there. Amidst these debates lies the relatively forgotten issue of the relationship between natural language and logical consequence – or formal logic at large. On the one hand, it might be thought that there are clear differences between natural languages and the formalisms of logic. For a starter, the grammar of natural languages is vastly more complex and intricate than the syntax encountered in most formal languages. More interestingly, however, is the apparent general consensus that also on the level of meaning, there is little congruence to be expected between logical constants or consequence relations as encountered in proof systems and models, and the meaning of pertinent expressions or inferences in natural language. In this vein, one recalls Strawson's famous quip: "[...] ordinary language has no exact logic." (Strawson 1950, 344). On the other hand, the advent of formal semantics with its tools might give rise to the impression that when doing natural language semantics, we are really just studying one language among many, using the same tools as we do when studying formal ones. As such, it might be thought that the relationship between natural languages and the ones employed in formal semantics – chiefly model-theoretic tools of formal logic – is extremely close (cf. Glanzberg 2015, 71).

Michael Glanzberg's paper *Logical Consequence and Natural Language* (2015) is a rare exception among contemporary authors, studying the relationship between natural language and formal logic, specifically logical consequence. His overall conclusion is that natural language and logic come apart, in the sense that natural language neither comes pre-equipped with a distinguished set of logical constants, nor does it ultimately generate a consequence relation (cf. ibid., 71ff. & 97).

Such a mismatch between logic and natural language would give rise to two puzzles, however. First, it may call into question our practice of using the tools of formal logic to study natural language meaning, as it is common in the enterprise of formal

[1] One can see the difference between a concept F and a conception *of* F to be that the former pertains to a piece of language as well as linguistic competence – encapsulated in the phrase "possessing the concept (of) F" – while the latter is a theory or account of 'F-ness' (cf. Künne 2003, 1f.).

semantics. For example, authors appeal to logical analyticity in explaining certain linguistic phenomena, such as why "believing whether" is ungrammatical (cf. Aloni et al. 2017). With what right do we appeal to logical notions such as L-analyticity, if natural language turns out to be devoid of them?

Second, such findings stand in deep tension with a basic observation about the epistemology and methodology of logic. Similarly as with the methodology of formal semantics, a mismatch between natural language and formal logic would give rise to a related concern: With what right do we rely on natural language examples when constructing, assessing and shaping logics? For example, it is commonly thought that the impetus for propositional logic is a formalisation of our reasoning with certain words, such as "and", "or", "If..., then...", "neither... nor..." *et cetera*.[2] However, the issue goes deeper. To give another example, when Hanfling discusses Russell's claim that the latter's analysis of existence claims can only proceed within a newly invented formal language,[3] Hanfling makes the following point:

> [I]f Russell really could not 'state the point correctly' in ordinary language, then how could we understand it? Is the statement in the new language [*viz.* $\exists x(man(x))$] translatable into the old? If it is, then the point can be stated in the old. If not, then we could not tell what the point is.
> (Hanfling 2000, 160)

Despite appearances, Hanfling is not repeating the paradox of analysis here,[4] but instead points towards a – for want of a better term – 'transcendental' observation concerning the relationship of natural language and logic. If a piece of formalism is not just supposed to be an exercise in algebra or formal syntax, but to be tied to formalising reasoning, it must be able to be put in an appropriate relation to

[2] One can take the cue from Stokhof (2007, 615) and assume that the best source for gaining insight into such commonsensical assumptions are textbooks on the matter. For an expression of the thought that formal logic is grounded in informal, ordinary reasoning – as manifested in natural language – cf. van Dalen (2013, 5f.), Hughes & Cresswell (1996, 4ff.) and Smullyan (1995, 4).

[3] Specifically, to analyse "Men exist." as "There is an x such that x is a man", cf. Hanfling (2000, 160).

[4] As a general proponent of conceptual analysis, this could hardly have been his point. As a proponent of so-called 'connective analysis' specifically, the paradox is, moreover, not an issue, since connective analysis tends to study the relations *between* concepts, not analyse a *single* concept into, according to the paradox, ultimately the same thing (cf. Beany 2021, sect. 6.4 & 6.8).

the medium of ordinary reasoning – natural language.[5] Most straightforwardly, this relation is one of *translation*, for otherwise, it would seem to remain unclear what the formalism is supposed to *tell* us. If these observations are correct – as I would submit – a mismatch between logic and natural language would be a strange occurrence indeed: If natural language is devoid of logical notions, then how does it come we seemingly cannot count something *as* logic, unless it can be put into relationship with ordinary reasoning as it is manifested in natural language?[6]

It seems, then, that mismatches between natural language and formal logic are not without issues. I do not intend to claim that these issues are necessarily insurmountable, but it ought to be clear that they give rise to a predicament: How could we reconcile a mismatch with the ubiquity of natural language examples in formal logic – both as a methodology *and* concerning our understanding of formal logic itself? And with what right could we continue to use notions from formal logic to study linguistic meaning?[7]

It is precisely with the start of these issues that the following investigation makes its contribution. If we had good reasons to suppose that such mismatches are either not deep enough to cause methodological or epistemological worries, or, even better, that such mismatches do not arise to begin with, we would be able to forestall these issues, or at least soften their impact. In what follows, I will argue that at there is no 'deep', *conceptual* mismatch between the notions of formal logic and natural language. In other words, the criteria for logical constanthood and consequences can, *in principle*, be satisfied in natural language. Moreover, I will offer reasons to be optimistic that there is, in fact, *no* such mismatch, i.e. we actually do find a consequence relation in natural language. The main issue with Glanzberg's arguments, or so I shall argue, is that they rely on problematic assumptions both about linguistic meaning and about the defining criteria of logical notions. I will demonstrate that by rejecting referentialist truth-conditional semantics and the corresponding model-theoretic conceptions in favour of *proof-theoretic* ones accompanied by *inferentialist*

[5] Hence "transcendental", in the sense of pertaining to the possibility of understanding something – in this case a piece of formal logic.

[6] "Ordinary" in both of Ryle's senses: 'ordinary' as pertaining to the 'average Joe and Jane', i.e. as opposed to specialist reasoning, and 'ordinary' as opposed to 'irregular' reasoning, this time including specialist reasoning under "ordinary" (cf. Ryle 1971, 301-304).

[7] Cf. Cotnoir (2018, 319ff.) for implications about formal semantics, and cf. Hjortland (2019, 260f.) for ramifications for the epistemology of logic, both specifically coming from Glanzberg (2015).

assumptions about meaning, we can avoid mismatches – at least in principle, if not completely.

To this end, my thesis is structured as follows. I will start off by considering Glanzberg's findings in detail (chapter 2), discussing all three of his arguments and unearthing his basic assumptions. Moreover, I will also briefly discuss some of the criticism that has been levelled against Glanzberg's arguments, preliminarily finding them wanting. Having dealt with the mismatch between referentialist truth-conditional semantics and model-theoretic conceptions, I will move on to introduce inferentialist metasemantics and semantics (chapter 3), both as a global account of natural language meaning and as a specific account about the meaning of logical constants as they appear in formal logic. Having set the stage, I will then consider proof-theoretic conceptions of logical constants and consequence, as they have been motivated and developed on inferentialist grounds (chapter 4). There, I will generate so-called 'realisability conditions', criteria which are individually necessary and jointly sufficient for the obtaining of the relevant notion in natural language. On the basis of said criteria, I will then assess to what degree these criteria mismatch with the inferentialist understanding of natural language, both on the conceptual level by comparing the realisability conditions with the meaning-theoretic assumptions of chapter 3, but also on the extensional level by considering a specific example of a purported logical constant (chapter 5). The overall conclusion, as I have indicated above, will be that there is no conceptual mismatch between natural language and formal logic, and that there are reasons to be optimistic about the actual existence of logical consequences in natural language. Lastly, I will revisit some of the larger themes from above and point towards further directions of research (chapter 6).

Before we proceed, a few remarks are in order. First, I am afraid my thesis will not make any *direct* contribution to the ongoing debate between inferentialism and its adversaries. Instead, the core contribution of the investigation is to study the parallel set-up to Glanzberg under use-theoretic lights, and see whether we get an analogous result. That said, the reader is free to agree with the epistemological and methodological observations concerning logic from above, and hence see the failure of referentialism and the success of inferentialism to account for these as an argument in favour of the latter, while putting the onus on the referentialist to produce an epistemology of logic that can account for the above considerations.

Second, I will not use, much less rely, on any specific definition of what counts as a natural language or a formal one. The reason is that it seems clear enough that the languages of logic are paradigmatic examples of formal languages, whereas English is undoubtedly a natural language. Thus, to the degree that any such distinction has currency, these languages will have to fall on their respective side of the distinction. Since these are the only languages we will be considering, we do not need any further, precise account of the distinction.

Third, as I have indicated above, there is a clear sense in which natural language and formal logic mismatch. It is obvious that the syntax of formal logic is vastly more streamlined than the grammar of natural languages, and thus we can find many discrepancies between the syntactic behaviour of logical constants as they appear in formal logic and their purported cousins in natural language. However, it is far from clear how such a mismatch might be of epistemological and methodological interest. Rather, it seems that the interesting comparison happens on the level of meaning, not grammar. As such, I will disregard such trivial types of mismatches, and focus on the semantical level.

Fourth, and last, for reasons of space I will have to confine myself to the study of sentential connectives, leaving the study of other kinds of logics for another day. Specifically, this means that I will not investigate issues related to quantifiers or modal operators. In the former case, this has the added benefit of sidestepping difficult issues concerning the logicality of higher-order quantification, which seems to occur with the use of "everything" in natural languages. After all, "Sallie has become everything she ever wanted to be." seems to contain a straightforward quantification over properties. Alas, we will have to sidestep this topic completely in what follows, and consider only the narrow class of sentential connectives.

2 The Glanzbergian Findings

2.1 Glanzberg's Model-Theoretic Mismatch

We now proceed to look at Glanzberg's arguments more closely. In his paper *Logical Consequence and Natural Language* (2015), he offers a rejection of the idea that the connection between formal logic and natural language is "[...] extremely close [...]" (Glanzberg 2015, 71), in the sense that, as logic studies various languages, natural language just happens to be one of them (ibid.). Specifically, the idea will be to show that natural language semantics in its contemporary form does not give rise to a relation of logical consequence within natural language, nor does it come pre-equipped with a distinguished set of logical constants (81, 97 & 114[8]). To this end, Glanzberg makes a series of assumptions, to which we turn first.

2.1.1 Background Assumptions

Glanzberg considers the consequence relation to be the 'core feature of a logic', which accordingly allows us to measure the closeness of natural language and formal logic (71). He takes such a consequence relation to be about preservation of designated values (73). More specifically, he follows other authors, such as Beall, Restall and Sagi (2019, sect. 1 & 2), in conceiving logical consequence to be both *necessary* and *formal*, the latter being cashed out in terms of logical constants (76f.). To put things into one slogan, logical consequence is the necessary preservation of designated values in virtue of the form of the involved sentences. Moreover, Glanzberg assumes a specifically *model-theoretic* view about both logical consequence (79) and logical constants (77 & 96). He takes this to be a harmless assumption, given that it is the *most* probable route to getting a consequence relation in natural language (79). Thus, for there to be such a relation, we would need a range of *models*, over which we could study the preservation of the designated values (cf. 88, for example). In a nutshell, a set of premises Γ logically implies a conclusion φ just in case every model that satisfies – usually: 'makes true' – every sentence in Γ also satisfies φ, where the extensions of all expression are varied maximally across models, with the exception of the logical constants. Using model-theoretic tools as well, the latter are distinguished by their topic-neutrality, which is accounted for by their extensions

[8]In what follows, I will omit repeated use of "ibid.". Naked page numbers always refer the relevant pages in Glanzberg (2015).

remaining invariant under arbitrary permutations of the domain of individuals (cf. 96).

The starting idea is that logical consequence as a relation between meaningful sentences tells us interesting things about such connections, and, conversely, the basis of such consequence relations is to be given by those sentences (74). This gives rise to the, according to Glanzberg, rather uncontroversial *logics in formal languages thesis*, which considers formal languages to be the means of studying consequence relations:

> Logical consequence relations are determined by formal languages, with syntactic and semantic structures appropriate to isolate those relations. (ibid.)

This thesis, in turn, motivates another one, which echoes the introductory idea about the closeness between formal logic and natural language, the *logic in natural language thesis* (LNLT):

> A natural language, as a structure with a syntax and a semantics, thereby determines a logical consequence relation.
> (75)

It is precisely this thesis that Glanzberg wishes to investigate.

As far as general philosophical viewpoints on logic itself are considered, Glanzberg often relies on the traditional idea that concerning logical consequence, there is a substantial matter of fact about what constitutes such a relation, and that there is an important underlying notion of logic in play. Moreover, the latter is taken to provide the former (77f.). He calls this the 'restrictive view', and under such a view, logical consequence is (roughly) first-order classical logic (79).[9] On the other hand, more 'permissive views' might still adhere to the substantiality of the notion, but interpret necessity and formality more broadly, which might yield logical pluralism and possibly non-classical logics (78f.).

[9] It is not entirely clear to me why precisely *first*-order logic results from these requirements. As MacFarlane (2017, sect. 5) notes, all higher-order quantifiers pass the test of permutation invariance, not just first-order ones. Hence, if logical constants form the 'backbone' of logical consequence relations, why do we not also consider relations involving relevant higher-order quantifiers? It might be that Glanzberg wishes to avoid the controversial topic of the logicality of second-order logic. In any case, neither Glanzberg's arguments nor the counters to them are somehow dependent on this, as we will see in due time.

An important piece of terminology is the distinction between logical consequence, entailments and implications. The latter are to be seen as the most general relation of support between sentences, such as in inductive reasoning or Gricean implicatures. Entailments, on the other hand, are the usual inclusions between sets of truth-conditions, i.e. P entails Q just in case the truth-conditions of P are included in those of Q, which gives us the conditional that whenever P is true, Q is true as well. Glanzberg notes that logical consequence is to be seen as narrower than entailments, and entailments as narrower than implications (80f.). This can be naturally read as set-inclusion, of course, but it might make sense to understand Glanzberg in a different light. Specifically, I suppose Glanzberg to mean that "narrower" ought to be understood rather in qualitative terms: There are vastly more ways for sentences to support each other in general (implications) than, e.g., in specifically analytic ways (lexical entailments), thus entailments are more *demanding* than (general) implications. *Apart* from those, logical consequence is an even *more* demanding notion, which obtains only under certain circumstances – the presence of a range of models and logical constants as individuated by invariance criteria (cf. above), which, as we will see, *lexical* entailments in *absolute* semantics do not require. Lexical entailments are entailments that rely on specific meanings, and in the context of absolute semantics, their contributions to truth-conditions are not relativised to any model, but are 'absolute' in a sense to be specified below. Yet lexical entailments are exactly what we will find in natural language, as Glanzberg will argue. As such, logical consequence relations are not a 'subset' or 'subspecies' of (lexical) entailments. They differ in what they require to obtain, and hence are distinct kinds of things, though naturally both belong to the vast class of implications.[10] Although, in general, whenever φ logically implies ψ, we have that the truth-conditions of φ are contained in those of ψ, otherwise it would be possible for φ to be true while ψ is not. As such, entailments in general might be seen as including logical consequence relations. But, in order to reiterate, *lexical* entailments as they appear in *absolute* semantics do not contain *logical* consequence relations, or so I take Glanzberg to claim.

A last important assumption that Glanzberg makes concerns the nature of linguistic meaning. He "[...] take[s] it for granted that truth conditions provide a central aspect of [what speakers know when they know what their sentences mean]." (85).

[10]Glanzberg speaks of 'subspecies' only when relating logical consequence to implication (cf. ibid.). However, I see no reason why entailments are not also a kind of implication, and nothing in Glanzberg (2015) would suggest that he would object to that.

This knowledge is seen as an aspect of speakers' linguistic competence, namely of what they come to grasp when they understand their sentences (ibid.). He rules out other approaches, such as inferentialism and other use theories, on the basis that they have not been successful on empirical grounds (fn. 16).[11] It is, of course, this footnote that provided the impetus for my thesis. As such, it is important to note that the choice of assuming model-theoretic conceptions of logical notions is no coincidence, given Glanzberg's commitment to truth-conditional semantics. I take it that this is the reason why he thinks such conceptions are the most plausible route towards the LNLT (cf. above). Both approaches seem to be indebted to the same tradition, namely Tarski's,[12] and make use of his model-theoretic tools and ideas. Hence, we might expect there to be no mismatch between doing logic and semantics in the sense of the introductory idea above.[13] This will motivate my choice of studying inferentialism together with proof-theoretic conceptions in section 2.3.

2.1.2 The Argument from Absolute Semantics

Glanzberg makes the observation that both standard approaches to truth-conditional semantics – Montagovian and Neo-Davidsonian – are engaged in providing so-called *absolute semantics* (89ff.). The idea here is, in contrast to the *historical* Montagovian approach, to forgo specifying meanings in terms of satisfaction in models, and instead deriving Tarskian T-sentences. The crucial point here is that the latter pro-

[11]Since Glanzberg does not detail his issues with use theories, I will not spend much time on the issue, either. Two general remarks are important. First, use theories are generally less developed than their referentialist cousins (cf. Steinberger & Murzi 2017, 214.). As such, it must be checked whether the lack of empirical success is due to a genuine failure to make correct predictions, or due to a lack of testing. Second, it is important to understand the explanatory ambitions of each theory correctly. As we will see in section 3.2, the *explanans* that inferentialists start with – the inferences judged to be valid by the linguistic community – forms the *explanadum* for the referentialist. Clearly, however, we would beg the question against the inferentialist if we accuse them of failing to account for entailment-judgements, since they form their point of departure.

[12]There are subtle interpretational issues here, as far as the role of varying domains is concerned. On the one hand, Sher (1991, 41-46) argues that despite Tarski not explicitly mentioning them, he must have intended them in his model theory, as some of his results would not obtain otherwise. Against this, Williamson (1999, fn. 5) contests that this interpretation of Tarski also ends up attributing basic logical mistakes to him, since Tarski's claim about material and formal consequence coinciding is false under the Sher's interpretation, but true if we reject her exegesis. In either case, it seems clear to me that Tarski has been instrumental in giving both model-theoretic semantics and model-theoretic conceptions of logical notions its impetus. To the extent that motivation is concerned, however, that is all I need.

[13]Glanzberg himself notes that providing truth conditions for sentences might deceptively look like providing models, which arguably gave rise to the LNLT. Cf. (85f.).

vides no range of models (88). Since both mainstream approaches to contemporary natural language semantics are absolute, neither provides the grounds for generating a logical consequence relation, understood in model-theoretic terms. For we simply lack any range of models over which we could study the preservation of truth-values, since all truth-conditions are absolute. As such, there appears to be no logical consequence relation in natural language (ibid.). Moreover, since this argument does not hinge on any restrictive or permissive views on logical consequence, as both require a variation of models, both views are susceptible to it (103). Glanzberg calls this the *argument from absolute semantics* (AAS).

Since it will be important later when discussing Sagi's criticism of Glanzberg (cf. section 2.2), I wish to remark on the reason as to why absolute semantics is preferable over model-theoretic semantics. Glanzberg invokes an argument due to Lepore (1983), and it should be instructive to discuss an example. Consider the sentence "Erin is happy", and ignore any temporal considerations for the moment. For this sentence, model-theoretic semantics will derive truth-conditions of the following form:

For any model $\mathcal{M}$: $\mathcal{M} \vDash$ "Erin is happy" iff "Erin"$^{\mathcal{M}} \in$ "happy"$^{\mathcal{M}}$.
(cf. ibid.)

However, according to these truth-conditions, a speaker could understand such a sentence without knowing who Erin is, or without possessing the concept of happiness. In contrast, absolute semantics will arrive at:

"Erin is happy" is true iff Erin is happy,

which is "[...] what we want[...]" (ibid.), since it involves, in a sense, Erin and the concept of happiness.

More precisely, Lepore assumes that formal semantics are in the business of also supplying a theory of *understanding* (Lepore 1983, 173). In this regard, model-theoretic semantics fail, because relativised truth-conditions do not allow an addressee of "Erin is happy" to discern anything about the meaning of this sentence, apart from the fact that, if true, the sentence claims that whatever is denoted by "Erin" falls under the extension of "happy" – whatever subset of the domain that is (cf. ibid., 177f.).[14] Lepore further argues that starting from model-theoretic seman-

[14]What is more, it looks like Lepore anticipated Glanzberg's arguments when he claims that this result is unsurprising, given that model theory is a theory of logical consequence, concerned with a range of models, as opposed to the absolute semantics of a theory of meaning. Cf. (ibid., 181).

tics, we cannot get to specifying absolute truth-conditions, since that would require singling out the actual world. This, however, is not possible according to Lepore, since model-theoretic semantics do not allow us to figure out which possible world is the actual one (ibid., 182ff.).[15] It must be remarked that the idea that the T-sentences and the corresponding 'absolute' lexical entries ought to be in the business of modelling understanding is not universally shared, however.[16]

In the meantime, I take this to provide both a crucial assumption on Glanzberg's part and an important *independent* motivation for adopting absolute semantics, that bears *no* relation to the discussion surrounding the LNLT. There is good reason to do absolute semantics – or so it seems. I will come back to this point when discussing Sagi's criticism in section 2.2.

Another important observation is that the AAS equally applies to intensional semantics, such as in Heim & Kratzer (1998).[17] It might be thought that the switch to intensional semantics, where our semantic values will involve possible worlds (Heim & Kratzer 1998, 303-309), will allow us to generate a logical consequence relation after all.[18] However, Glanzberg is well aware of this (cf. 87). What counts is that we

[15]This passage is a bit puzzling, as there is a difference between possible worlds and models, as far as intensional semantics are concerned (cf. below). However, I think Lepore's point can be made for models just as well: Given the range of models we are entertaining, how do we come to know which one is the one 'we live in'? Arguably, we need a lot of information to distinguish possible worlds – or models – and if model-theoretic semanticists wish to arrive at absolute semantics, they must show how we get to this knowledge with only relativised truth-conditions – certainly a daunting task. Moreover, absolute semantics simply sidesteps this issue by relying on the ordinary meaning of things, where singling out the actual world or model is unnecessary (cf. ibid., 184).

[16]For example, Yalcin (2017) argues that in T-sentences and lexical entries as we find them in formal semantics, the right-hand side featuring the formal expressions supplied by the theory ought not to be seen as translations of the left-hand side, which features the natural language expressions (cf. ibid., sect. 12.3). His chief reason to think so is the fact that if they were translations, then we would be robbed of our explanatory ambitions, since the *explanans* – the right-hand side – would involve the *explanandum* – the left-had side (cf. ibid., 344ff.). If that were so, Lepore's argument could not go through anymore, for if we reject the disquotational reading of the relevant entries, we clearly could not supply a theory of understanding to begin with. There are two main issues with this approach, however. First, it is not so clear if we really rob ourselves of *all* explanatory ambitions simply because we allow for disquotational readings (cf. Glanzberg 2014, for a more fine-grained approach to the issue). Second, even if such an approach to semantics would allow a reintroduction of some models back into semantic theorising (cf. Yalcin 2017, 343f.), it is not clear if we would get a sufficiently *large* range of models to generate a *logical*, hence *formal*, consequence relation. I will return to this point when discussing Sagi in section 2.2.

[17]I chose type-theoretic semantics since I am more familiar with them myself. However, I am confident nothing hinges on this choice, as should become clear during my argument.

[18]This has been proposed by many of my fellow students and by staff members at the ILLC, when discussing Glanzberg's arguments.

get a range of *models* over which we can trace relevant preservations, which is "[...] *built into* the basic apparatus of semantic theory." (88, my emphasis). This is not the case here. The Carnapian intension of "happy" is some function from possible worlds to extensions in those worlds, specifying for each world the things which are happy (cf. Heim & Kratzer 1998, 305). All of this still takes place *within* absolute semantics, though: We do not vary the meaning of "happy" when considering different possible worlds, holding the meanings of some specified set of logical constants fixed. We hold *this* meaning constant from the start. Thus, I do not think that intensional semantics are a way out of the problem. For there is still no variation of models provided by the semantic theory *itself*.

Faced with the AAS, a proponent of the LNLT might look for other sources of logicality in natural language. Glanzberg anticipates this, and his *argument from lexical entailments* (ALE) and *argument from logical constants* (ALC) are supposed to block such moves.

2.1.3 The Argument from Lexical Entailments

While we may not get a straightforward consequence relation out of natural language semantics, it might still be contested that the various entailments we find in natural language provide a basis from which to generate one (91f.). After all, truth-conditional semantics give each sentence truth conditions, which in turn give us entailment relations. Moreover, such entailments are an ubiquitous phenomenon in natural language, and any semantic theory needs to account for them (93f.).

As already seen by the condition on formality, though, such entailment relations themselves cannot be taken as logical consequences. The entailment of "John is not married" by "John is a bachelor" depends on the specific meaning of "bachelor", which is not a logical constant by restrictive lights (92). Even if we were to start treating such items as logical constants under more permissive considerations, we face two general difficulties. On the one hand, almost every word in natural language triggers some entailment effects, hence our class of logical constants will be huge. This would trivialise the formality constraint (94). On the other hand, the model-theoretic criterion of permutation invariance for logical constants fails for these words (96). Hence, Glanzberg concludes, entailments are *lexical*, not formal, and thus cannot give rise to a consequence relation either (96f.).

A potential hope for more permissive views would be to suggest that lexical entailments are generated by a (sufficiently) small and homogeneous group of features, which could then be treated as logical constants (105). At an earlier point, Glanzberg considers the idea that classes of lexical entailments are generated by a common structure, and hence are not as heterogeneous as they may appear at first glance. This common structure might then be the grounds on which expressions figuring in lexical entailments might still be considered logical constants, given that they share common features and hence may not undercut formality (95f.). However, Glanzberg thinks it is doubtful that such features are forthcoming, since there are currently no candidates to be found, and the two obstacles faced above might anyway remain genuine issues. As such, permissive views face a challenge here, at least (cf. 105f.).

2.1.4 The Argument from Logical Constants

The third and last approach to saving the LNLT as discussed by Glanzberg argues that natural language at least provides us with a distinct class of logical constants, which can then be used to generate an appropriate consequence relation. Glanzberg blocks this line of argumentation by pointing to two things: (i) The absolute semantics of natural language do not endow the logical constants in natural language with a consequence relation and (ii) natural language itself does not even distinguish between logical and non-logical expressions (97).

Let us look at (i) first. The example under consideration is the quantifier "most", with its standard truth conditions via the theory of generalised quantifiers:

$$[\![\text{“most”}]\!](A,B) \qquad \text{iff} \qquad |A \cap B| > |A - B|.$$

(cf. 98)

As Glanzberg observes, these truth conditions do not contain any quantification over models, not even tacitly (ibid.), which is already recognised by generalised quantifier theory in its distinction between local and global quantifiers. For example, the global version of "most" would be a function which assigns to each model $\mathcal{M}$ the extension of $[\![\text{“most”}]\!]^{\mathcal{M}}$, which is just the set of pairs (A, B) satisfying the truth conditions above. Absolute semantics do, and only *can*, deal with the local versions, as absoluteness bars any quantification over models, yet we would need the global versions to distinguish such quantifiers as logical constants, given the invariance criterion. Additionally, only the global versions could generate genuine consequence relations. For only they involve the variation of domains – without such, any entailment generated by such expressions could, in the context of absolute semantics, only ever be lexical (98f.). Since this part of his argument relies once more on absolute semantics, it applies equally well to permissive views (106).

For convenience's sake later on, we formalise this argument in its general form. I will come back to this part of the ALC specifically in chapter 5, when we discuss the inferentialist response. Meanwhile, we notice that the ALC has the following form:

(P1) In order for any purported logical constant in natural language to generate a logical consequence relation, such a constant would have to pass the invariance test, hence would have to be identified as *logical* to begin with, *and* allow the tracing of truth-preservation in its associated entailments over a range of models.

(P2) In order for a logical constant to do so, natural language semantics would need to involve relativised truth-conditions, and hence could not be absolute.

(P3) However, natural language semantics *are* absolute, as they could otherwise not account for linguistic understanding, which they should.

(C1) No logical constant could be identified as such, and whatever entailments the expression in question generates could only ever be lexical, i.e. cannot involve a range of models, but instead provides entailments only in the context of absolute semantics.
(From (P3) and (P2).)

C2) No purported logical constant in natural language could generate a logical consequence relation.
(From (C1) and (P1).)

As we can see from this presentation, the issue lies once more – as it was the case with the AAS – in the mismatch between the requirements for logical notions, as imposed by model-theoretic conceptions, on the one hand, and the requirements for semantics in order to do its explanatory work, on the other.

As for (ii), Glanzberg asks whether there are any semantic or grammatical properties that would distinguish logical constants from others, and concludes there are none. He notes that logical constants belong to the class of so-called *functional* expressions, which includes words such as "that" or comparative morphemes such as "-er". This family is thus rather big, so treating all of them as logical constants runs into the danger of trivialising formality once more (101). Additionally, the grammatical properties of this class make no distinction between the logical and non-logical expressions therein. Hence, according to Glanzberg, there does not seem to be any grammatical or semantic property[19] shared by the logical constants which is not also shared by other functional expressions. As such, natural language does not come pre-equipped with a characterisation of logicality (102).

On a more permissive approach, it might be possible to consider all functional expressions to be logical constants. Glanzberg thinks that it is doubtful that this will result in a single, coherent consequence relation, though. As such, at least one part of the ALC applies in full force to permissive views as well, while the other raises issues for the view (106).

2.1.5 Summary

These three arguments, then – AAS, ALE and ALC – constitute Glanzberg's attack on the idea that natural language, considered as a syntax with a semantics, gives rise to a logical consequence relation.[20] At this point, we make two further general observations. First, his arguments ultimately rely on three key assumptions, as should have become clear above:

(I) **Truth-conditional Semantics**: Truth-conditions provide a central aspect of semantic competence, and, accordingly, ought to be a key ingredient of studying meaning.

(II) **Absolute Semantics**: Natural language semantics are currently done relying on absolute truth-conditions, as opposed to the relativised notion of model-theoretic semantics. Moreover, this choice is deliberate and justified – relativised truth-conditions cannot account for linguistic understanding, while

[19] Recall that *in the context* of absolute semantics, classic logical constants are not distinguished by permutation invariance. The entailments they give rise to are semantically on par with other lexical entailments, as (i) suggests.

[20] This interpretation of the arguments also seems to coincide with Glanzberg's own understanding of them, cf. Glanzberg (2020, 10f.).

absolute ones do. This last bit rests on the assumption that formal semantics is also in the business of supplying a theory of linguistic understanding.

(III) **Model-theoretic Conceptions**: Both logical consequence and logical constants are to be analysed in model-theoretic terms.

Given this list of assumptions, a natural question to ask is whether any of them could be given up, and for what reason. Concerning those who accept (I), there have been dissenting voices about the status of (II), prominently among them Sagi (2020). I will discuss the issue below. Other than that, the question whether a rejection of (I) (or (III)), and the resulting rejection of the other,[21] gives rise to similar issues remains an open question – precisely the one I wish to tackle in this thesis.

Now, concerning this task of investigating the analogous case of proof theory and inferentialism, the second observation is Glanzberg's overall methodology. The idea is to study what the given conceptions of logical notions require in order for them to obtain (henceforth: 'realisability conditions') and then see if natural language actually meets these requirements. This will arguably depend on the philosophical assumptions about the nature of language – as noted before, there is not much chance that inferentialism about linguistic meaning in general is going to match well with model-theoretic conceptions of logical consequence. With this blueprint in hand, I will proceed similarly down below when investigating proof-theoretic conceptions of logical consequence.

Before turning to such investigations, I wish to discuss the potential rejection of assumption (II). If arguments against (II) would prove sound, it could then be claimed that formal semantics can make use of a space of models after all. Of course, this does not invalidate checking the potential predicament of proof theory and inferentialism, but it would undermine the 'big picture idea' that referentialism has a harder time accounting for our initial methodological and epistemological observations in relation to formal logic.

2.2 Interlude: Against Absolute Semantics

As I said before, the primary aim of my thesis is to look into the analogous case of the Glanzbergian predicament for proof theory and inferentialism, not to settle the

[21] After all, it is unclear what an inferentialist in semantics is supposed to do with model-theoretic conceptions of logical consequence, for example.

issue surrounding the status of absolute semantics. Since it is not my primary focus, I will focus on only one source – Gil Sagi's *Considerations on Logical Consequence and Natural Language* (2020, *forthcoming*), as Glanzberg's only direct critic so far. I will survey some of her responses, and tentatively conclude that her strategy cannot work without ultimately begging the question against Glanzberg.

Let us start with the AAS. Sagi counters it in two ways: by denying that contemporary natural language semantics are always absolute, and by arguing that, in either case, Glanzberg is guilty of begging the question against the LNLT. The reason why contemporary semantics are not always absolute is due to considerations stemming from Zimmermann (1999). The latter claims that models serve the important function of modelling the linguist's ignorance concerning the precise extensions of expressions in natural language (Sagi 2020, 8). As she further argues, this is not just an ignorance on the linguist's part, but also on the speaker's part, given that knowledge of the precise extensions of various terms is not always an ingredient in linguistic competence. After all, there is no reason to suppose that in order to understand the word "smokes", we need to be able to list everyone that smokes – nor could we, realistically. This claim serves to forestall the objection that these models are merely a result of less than ideal theorising, hence do not represent anything real about the phenomenon. For if such ignorance can sometimes not be overcome and is common for competent speakers themselves, the resulting models ought to be seen as representing something real about linguistic competence after all, namely precisely this kind of ignorance. Thus, they are not mere 'artefacts' of the modelling process (ibid.).

As I have indicated earlier when mentioning Yalcin (cf. fn. 16), it is less clear whether this allows us to introduce enough models to generate a proper *logical* consequence relation, which would be needed to properly account for the formality constraint on logical consequence (cf. Beall, Restall & Sagi 2019, sect. 1). Consider the predicate "x is a hair on Simon's head". Now, making certain harmless mereological assumptions, the cardinality of the extension of said predicate ranges somewhere between 0 and n, with n being the number of particles in the universe. Since no one – much less I myself – knows the precise extension of said predicate, it seems that we, following Sagi, need to introduce the corresponding number of models in order to account for this ignorance. The issue, however, is this. If the motivation for reintroducing the models was to account for this ignorance, we cannot add more models than that. But then, "The number of hairs on Simon's head is between 0 and n" becomes a straightforward *logical* truth, and hence a logical *consequence* of any set of premisses, because the models that natural language semantics provides us with cannot falsify it. However, such empirical, contingent matters that do not involve the form of sentences cannot be logical truths or consequences.

Let us, then, turn to the issue of Glanzberg begging the question. As Sagi argues, if natural language semantics are absolute, and this absoluteness bars us from generating a logical consequence relation, all this establishes is that

> [...] the study of truth conditions in natural language is not identical to the study of logical consequence in natural language, a mark of the difference is that one uses a range of models and the other does not. Glanzberg begs the question when he looks for logical consequence in natural language by looking at a discipline which he defines through its subject matter, which is not logical consequence.
> (Sagi 2020, 9)

Moreover, she understands Glanzberg as claiming that every logical consequence relation is an entailment, and since entailments do feature in natural language, all we need to do in order to study logical consequence in natural language is to adjust our tool-kit accordingly. Additionally, the framework of formal semantics already gives us the grounds on which we can introduce the appropriate range of models, which requires less information than absolute semantics, since, for example, the denotation of proper names can just be any individual of any domain (ibid., 9f.).[22]

I do not think Glanzberg would disagree with Sagi's claims to quite some extent. After all, section 3.5 in Glanzberg (2015) details how we get to such a relation, and indeed, this does include a reintroduction of models (Glanzberg 2015, 108f.). In other words, it seems to me that both agree that we *can* get a consequence relation on the basis of the framework of absolute semantics, namely by imbuing it with a range of models. So where is the issue? Glanzberg sees this reintroduction as going beyond what speakers know when they know the meanings of their words, while Sagi thinks this is nevertheless revealing about the logical structure of natural language. As such, in order for there to be genuine disagreement, it seems to me that Sagi is committed to the position that speakers nonetheless have a degree of competence with some logical notions, such as logical constants and certain entailments (cf. Sagi

[22] In this regard, recall Lepore's discussion about singling out the actual world in fn. 15.

2020, 9f.).

This raises a question, however. How did they receive such competence? If it were to arise from them being speakers of a natural language, Sagi's accusation can no longer stand. As Hjortland points out, Glanzberg's musings seem to suggest that semantic intuitions cannot feature in the epistemology of logic (Hjortland 2019, 160f.). In other words, Glanzberg's conclusion that natural language and logic come apart renders the idea that our semantic competence with certain expressions is the basis of our knowledge of logic shaky, as I have indicated in the introduction. For this reason, speakers cannot be taken to have competence with logical notions, *in virtue of their possession of a natural language*, at least not without further motivation. Thus, in order for there to be genuine disagreement between Sagi and Glanzberg, Sagi must be committed to the idea that there are logical constants or consequence relations in natural language already, hence the need to adjust our tool-kit. But then, I must conclude that Sagi would be begging the question, if anyone is. If Glanzberg's conclusion gives us reasons to doubt that competence with logical notions could arise in virtue of linguistic competence, it is unfair to assume that speakers, *qua* being speakers, have such competence nonetheless, which would in turn justify the introduction of models precisely to study these features of language. If, on the other hand, Sagi thinks that such competence does not arise from being a speaker, she owes us an explanation of where it comes from, and why it justifies the reintroduction of the required models into the framework of natural *language* semantics. This also harkens back to our discussion of Lepore. As I remarked there, Glanzberg's motivation for adopting absolute semantics is independent of the issue surrounding LNLT. As such, it seems puzzling to claim that Glanzberg is begging the question by defining the subject matter of semantics in such a way as to exclude the possibility for logical consequence.

In fact, at this point, the question arises whether any kind of linguistic phenomenon *could* motivate the introduction of a sufficiently large range of models into semantic theorising, *apart* from *logical* notions. I would be inclined to say "no". Yet if that is so, we should not be surprised we get the impression that people beg the question in this debate. Recall: We were looking for a motivation to introduce enough models to study logical consequence in natural language. Now, if it turns out the *only* thing that would give us enough models to study such a relation is the relation itself, it becomes clear that the debate surrounding the LNLT, at least within the truth-conditional-referentialist camp, cannot proceed solely by considering the status of absolute semantics. In other words, it might be that in order to make progress, the debate needs to look for further reasons as to why natural language could, or could not, contain such a relation.

With this lesson in mind, I would claim that her other arguments fail on similar grounds. For example, her insistence that determiners such as "most" *do* satisfy invariance criteria is naturally shared by Glanzberg (cf. Glanzberg 2015, 110).[23] The latter's point is that natural language semantics itself does not have the necessary ingredients to check those criteria in virtue of also being in the business of supplying a theory of understanding. In other words, unless we are given a good reason why a sufficiently large range of models needs to be introduced into the framework of formal semantics, this line of reasoning cannot be made to work against Glanzberg, not without tacitly assuming speakers to be competent with logic *qua* being speakers of a natural language, something that Glanzberg's conclusions seem to question.

Of course, more would need to be said. Sagi makes further arguments, drawing from work done by other authors and herself on other occasions (cf. Sagi 2020, sect. 4 & 5). However, as I was hopefully able to show, her main counters to the above arguments are in need of further development. Specifically, the fundamental question would be with what *right* we could introduce a sufficiently large range of models back into the framework of semantic theorising, apart from already assuming that speakers are competent with logical notions on the basis of their competence with a natural language. For otherwise, *we* would beg the question against Glanzberg.

Luckily for us, neither inferentialism nor proof-theoretic conceptions of logical consequence make central use of truth-conditions, much less models. In fact, this discussion might point towards an issue in *referentialist* theorising overall.

[23] I cannot delve too much into her various counters due to reasons of space. Suffice to say, I do not think that her main argument against the ALE works either, for the simple reason that she reads Glanzberg as claiming that entailments include logical consequences. If that were the case, and we need to model speakers' competences with entailments, it might seem that we also need to account for their competence with logical consequences (cf. Sagi 2020, 12f.). However, as I have argued in section 2.1.1, there is little reason to understand Glanzberg as claiming that logical consequence is a subspecies of entailments. Hence, without that assumption, it seems her counter is hardly convincing, for then, the lexical entailments of natural language no longer include logical consequence relations.

2.3 An Alternative Emerges

As we have seen from the argumentation so far, the project of truth-conditional formal semantics seems to lead to the conclusion that natural language neither contains a distinguished set of logical constants, nor does it contain logical consequence relations. Glanzberg's arguments to this extent are based on three assumptions: truth-conditional semantics, absolute semantics and model-theoretic conceptions of logical notions. It is important to notice that the project of formal semantics as it is presented above is a deeply *referentialist* project: The key aspect of linguistic competence are truth-conditions, and these are generated via compositionality of constitutive expressions, in the form of extensions, referents and satisfaction.

The central point of contention was specifically the use of *models* in capturing these relations of satisfaction, reference and extension. The same notions also give rise to the idea that logical constants are demarcated by relevant model-theoretic tools, and that logical consequence is essentially preservation of designated values over a sufficiently large range of models. Seeing how involved and open these issues are, one might wonder how the LNLT fares with respect to *other*, fundamentally *different* approaches to linguistic meaning and logical notions. If we were to rely on *proof theory* instead of model theory, i.e. the other big pillar in contemporary formal logic, what kind of characterisations of logical notions do we get? What would be the accompanying view on linguistic meaning? And most importantly: Do we get a mismatch between those two as well?

The answers to these questions will be given in the following chapters. We begin by surveying and motivating so-called *use theories* of linguistic meaning, specifically inferentialism over referentialist approaches, and then see to what kind of conceptions of logical notions in proof theory they have given rise to.

3 Inferentialism

In this chapter, I wish to give an overview over the topic of inferentialism, specifically its motivation, overall structure and variants. This chapter is not designed to convince the devotee of referentialist truth-conditional semantics to change their convictions, nor would I expect it to. Rather, the goal is to showcase inferentialism as a well-motivated and serious alternative to mainstream approaches to linguistic meaning.

I will start by giving the 'canonical' arguments against referentialist theories, which in turn ought to motivate a use-based approach to explicating linguistic meaning. Next, I will explicate the notion of inferentialism and consider its different variants, over metasemantics to semantics and back. The choice on which type of inferentialism we will be focussing on will be made via motivated choices along the way. After that, I will close off by giving standard considerations as to why specifically *inferences* ought to be given central importance in understanding meaning, especially when it comes to distinctly *logical* vocabulary, which is what we will be concerned with for the rest of this investigation. To this end, I will first provide a general argument in favour of inferentialism, before paying special attention to *logical* inferentialism.

3.1 The Motivation for Use Theories

As far as analytic philosophy is concerned, use theories are a more recent invention. The dominant strand of thinking about linguistic meaning has undoubtedly been referentialist in nature during the first half of the 20th century, and it is thus no surprise that use theories found their early *raison d'être* in opposing and criticising referentialist truth-conditional approaches. As such, we will naturally encounter arguments that serve the dual purpose of motivating use theories while simultaneously weakening the case for referentialism. To give a classic example of the aforementioned historical trends, the early Wittgenstein in the *Tractatus* (1921) had clearly truth-conditional referentialist ambitions, with its focus on truth-conditions and the picture theory of propositions (cf. TLP, §2-4.024), whereas the later Wittgenstein spends a sizeable amount of the *Philosophical Investigations* (1953) dealing with his earlier conception, scrutinising and criticising it (cf. PI, §1-43). While our investigation is not historical in nature, we are nonetheless in a similar situation, coming from the previous chapter. The assumptions (I) to (III) left us with an unsatisfying picture, and so we, too, are looking for alternatives, especially with regards to jettisoning (I)-(III) altogether.

So what, then, are the reasons to doubt the referentialist picture of meaning, according to which the basic model of meaning is that of reference of words to objects in the world? Glock (2018) gives us two immediate challenges to the referentialist picture. First, there is an abundance of meaningful words in natural languages for which it is hard if not impossible to make sense of in terms of reference. The referentialist model is based on proper names, mass nouns and sortal nouns, where it admittedly yields plausible results, yet has a much harder time making sense of other grammatical categories, such as prepositions and exclamations (ibid., 65). For example, consider the exclamation "Ouch!", and ask yourself what kind of object it is supposed to stand for. It should be clear that "Ouch!" is not making any assertion when used, nor does it refer to anything (cf. PI, §27). In its ordinary use, it does *convey* that somebody is in pain, but it does so by means of *expressing* pain, not by 'pointing', as it were, to one's pain, much less by making an assertion about one's physical or mental state. If it were making an assertion, we should be able to legitimately reject or affirm its content. But this is clearly not the case: "Ouch!" - "Yeah, you are right!" is an exceptionally hard-to-contextualise[24] piece of dialogue. Moreover, it is hard to understand how an act of mere reference to an emotion[25] can succeed in *expressing* an emotion. If "Ouch!" were to function like an indexical referring to the utterer's (physical) pain (wherever it may be), then any description fixing the location of the pain would be co-extensional with it, and should in principle be able to express the speaker's pain as well. Yet clearly, while "Ouch!" may succeed in expressing the pain in my thumb, for example after hitting it by accident with a hammer, "The pain in my thumb!" is an awkward utterance at best, and would in any case not succeed in expressing pain *unless* accompanied by shouting or other pain-related behaviour. In that case, however, it is clear that the work of expressing pain would be done by the behaviour, not the phrase "the pain in my thumb".

Second, even in case of those expressions that *do* refer to something, their mean-

[24]In the sense of being able to describe a context or situation where, in this case, two speakers would have an intelligible conversation involving the exchange under consideration.

[25]For the more general worry concerning emotions and reference, cf. once again PI, e.g. §293, and Glock (2020, 205f.). The latter provides an especially succinct and modern formulation of the so-called private language argument, which ought to show that the idea of essentially private *mental* entities is conceptually incoherent.

ing is not the object they stand for. Glock uses an example due to Wittgenstein (PI, §40). Wittgenstein remarks that it is important not to confuse the bearer of a name with the meaning of a name. If someone dies, we say that the bearer of the name has died, not its meaning, and it is moreover intelligible to say things such as "Mr. X has died.", which could not be the case if the meaning of a name were to be identified with its bearer. For if they were the same, the name should become empty, rendering the aforementioned death report meaningless (cf. Glock 2018, 65f.). Similarly as in the discussion of the argument from absolute semantics above, one might think that a switch to intensional semantics might be able to fix the issue. One might argue that the intension of a proper name – usually a function from possible worlds to some entity (cf. Heim & Kratzer 1998, 304) – might change in its course of values when the bearer dies in the actual world, but the function itself does not vanish as such, hence the name retains *some* meaning. This solution will depend heavily on the details of the quantified modal logic used to provide such intensions, such as whether proper names are rigid designators and whether we allow for a free logic.[26] In any case, this cannot straightforwardly serve the referentialist, as it will potentially only compound further issues. For example, one can fall in love and marry the bearer of a name, but clearly not its meaning, and especially not a function of type $< s, t >$ (cf. Glock 2018, 66).[27]

The preceding two arguments are targeting what Glock calls '"Fido"-Fido' conceptions of meaning, i.e. those theories that take the relation of reference as the basic model of meaning. This by itself does not motivate use theories *per se*, as nothing in their defence has been said so far. Moreover, it does not disqualify truth-conditional approaches yet, as there can be truth-conditional semantics that are not inherently referentialist, such Davidson's truth-conditional approach (cf. Brandom 2000, 52, and Davidson 1967). As such, we need more arguments in order to motivate use theories.

Luckily for the use theorist, there are considerations directly in favour of the approach. A special status has the so-called 'constitutive argument', which serves both as a criticism of referentialist approaches and as a motivation for use theories.

[26] Cf. Gamut 1991, chap. 3, for an overview.

[27] Provided, of course, the advocate of the intensional treatment is still committed to treating the meaning of a name to be its bearer. Then again, the corresponding function is usually taken to be constant, with the bearer figuring as the only value throughout. Thus, there might not be too much difference between the intensionalist and extensionalist treatment as far as the aforementioned identification of the bearer with the meaning is concerned.

The argument starts with the mundane observation that meaning is something that can be explained or taught to somebody. However, straightforward attempts at conveying the meaning of "yellow", for example, cannot succeed, *unless* the addressee *already* understands a language. Consider the explanation "'Yellow' designates the property yellow/the class of yellow things/etc.". This is clearly intelligible *only if* the addressee already speaks English *and* understands the word "yellow". Even if we start to mix languages, the point remains: "'Yellow' ist das englische Wort für Gelb." is only intelligible if you understand a sufficient amount of German *and* the German word "Gelb" (cf. Skorupski 2017, 74). Now, attending to the way the meaning of a word is actually taught to someone who does not have another language to fall back on, we see that this is, in general, done by teaching *rules* for the *correct use* of a word.[28] These rules must thus constitute knowledge of meaning, if they are what somebody needs to be able to follow in order to be considered as knowing what a word means (cf. ibid., 74f., and Glock 2018, 66). Glock makes the further comparison between language acquisition and learning how to play chess. The latter is not done by associating the pieces' names with objects, but by learning how to move them on the board in accordance with the rules of the game. In the same sense, we learn the meaning of a word (cf. Glock 2018, 66).

At this stage, an important caveat is in order. As I have labelled it earlier, this argument ought to be both a criticism of referentialism *and* an argument in favour of use theories. I hope that the positive aspect is clear enough: Use theories are able to capture the way that meaning is actually acquired, as opposed to the referentialist, who faces a challenge in explaining language acquisition according to their model. However, and this is the caveat, as Skorupski points out, this argument is concerned with *metalinguistic knowledge*, not the designations of expressions *per se* (cf. Skorupski 2017, 75). To the degree that a referentialist insists that their model of meaning is about what speakers know when they know what their words mean, it seems to me the argument *does* strike, as it would become mysterious how speakers

[28]As Paul Dekker told me in private conversation, there is a subtle distinction here. What is most plausibly taught is the correct use itself, resulting in tacit knowledge-how of applying the relevant expression. On the other hand, we can – and sometimes do – teach the rules themselves, in the style of "If this-and-this is the case, then you can apply "F" to this-and-that.". However, I do not think the distinction matters much for what I have to say. For even if inferentialist semantics ultimately deal with the rules themselves – hence seemingly suggesting that knowledge of meaning is *explicit* knowledge of the pertinent rules – it seems, nonetheless, possible to read lexical entries as merely claiming that these rules are operative in knowing the meaning of an expression – be it explicitly or tacitly.

manage to acquire this knowledge, then. However, this argument *alone* does not invalidate the idea that the *semantic values* (as Skorupski puts it) of words cannot, at least on occasion, be what the referentialist takes them to be, namely objects, classes of such, etc. Unfortunately, though, Skorupski leaves the relation of meaning to semantic value open. If what we are interested in is a theory of meaning, the shift or introduction of some further technical notion seems dodgy. For if we wish to stay on the topic of meaning, we might do better not to introduce further notions, lest we wish to invite the challenge of further having to establish the philosophical relevance of such notions in relation to meaning (cf. Glock 2017, 90). Indeed, if semantic values are referents, classes and the like, yet those cannot be taught or explained to someone without presupposing linguistic competence on their part, then the conclusion seems to be that semantic values are not meanings. For the latter *can* be taught or explained, even to someone who has no other language to fall back on.

Let us take stock. We have seen three arguments, two of which exclusively target referentialist models of meaning, one of which targets the latter to the extent it entails a referentialist understanding of metalinguistic knowledge and language acquisition, and in any case motivates a use-theoretic approach. Together, they form a coherent front against specifically referentialist theories of meaning *in general*, but, as noted earlier, not necessarily against truth-conditional ones, such as Davidson's. We have to take a further step to motivate use theories over truth-conditional ones. Moreover, at this stage, the term "use" is not specified at all, and involves a large class of activities. Both of these issues will now be taken care of by reviewing specifically *inferentialist* theories of meaning, which, at least as far as the *declarative part* of a language is concerned, identify the relevant notion of 'use' with the drawing of inferences.

3.2 Varieties of Inferentialism

Before answering the question of why one ought to be an inferentialist, it ought to be instructive to get a better grasp of what inferentialism actually entails. Steinberger and Murzi characterise it as follows:

> [Inferentialist theories form] the particular class of use theories that gives *inference* pride of place in their account of meaning.
> (Steinberger & Murzi 2017, 197)

The reason why a restriction to a specific type of use is desirable comes from the observation that, plausibly, not all aspects of use are equally central to a word's meaning. Inferentialism, as a subspecies of conceptual-role semantics, which hold that the content of symbols is determined by their role or use in thought or speech, restricts the semantically relevant aspects of a word's meaning to its role in the regularities and conformities with extant inferential practices, or some subset of those (cf. ibid., 197f.). As a spoiler for what is to come, the classic example of an inferential characterisation of meaning concerns logical vocabulary. In relation to the word-pair "if"-"then", the inferentialist take would be to see their meaning as being constituted by their inference rules, such as *modus ponens* and *conditional proof*. It is a general characteristic of inferentialism that its mode of explanation starts with the meaning of declarative sentences, and then moves from these to the meaning of their constituent expressions in a top-down direction of explanation, as opposed to the bottom-up approach of referentialist theories (cf. ibid., 197 & 200f.).

Before we deal with logical inferentialism, we can get a better grasp of inferentialism by considering its different variants, and how they fit together. I will focus on the kinds of inferentialism relevant to our investigation, and ignore others, both in the interest of space and succinctness.

Given the kind of investigation we are after, we are interested in the varieties of inferentialism that serve as a theory of linguistic meaning, specifically those of the public natural languages. As such, we will not entertain individualistic interpretations of inferentialism (cf. ibid, 202f.). Moreover, given that our primary focus will be on proof-theoretic views of logical consequence, which are in turn advocated by authors such as Dummett (1991), and given that we will consider inferentialist theories that go beyond logical vocabulary by encompassing entire languages, such as Brandom (1994, 2000) and Peregrin (2014), I will follow these authors in understanding conceptual content to be at least going hand-in-hand with mastery of public linguistic meaning, if not even being reliant upon the latter (cf. Steinberger & Murzi 2017, 198f.). This results in the general idea that conceptual content can only be accounted for via an account of linguistic meaning, in the sense that linguistic competence is required in order to ascribe conceptual competence to a creature.[29]

[29]Given my personal sympathy for the private language argument (PLA), I am more than happy to accept this conclusion anyway. The meaning of "possessing the concept F", must, on pain of conceptual incoherence – given the PLA – partially consist in public, observable behaviour. Given that concepts are once more things that can be explained and taught via language, the public behaviour partially constitutive of the meaning of "possessing the concept F" arguably precisely

With this qualification out of the way, we arrive at our first real decision point. There is a general distinction between semantics and metasemantics. The former is taken to be occupied with studying the actual meaning of expressions, whereas the latter is said to study the foundations of the former, such as on what grounds expressions have the meanings that they have, or what constitutes knowledge of such meaning (cf. Burgess & Sherman 2014, 1f., and Steinberger & Murzi 2017, 199). With respect to this distinction, inferentialism in its most 'general' form is a metasemantical thesis about the order of explanation between the meaning of an expression and its use in language, relating either to the explanatory question of how expressions get the meaning they have, or the epistemological question of what it takes to know the meaning of an expression. Accordingly, inferentialism on the level of metasemantics can be seen as giving rise to two theses (cf. Steinberger & Murzi 2017, 199):

1. The meaning of a linguistic expression is determined by its role in inference (explanatory).
2. To understand a linguistic expression is to understand its role in inference (epistemological).

Given the constitutive argument above, we are already committed to the epistemological claim. We will become committed to the explanatory one in section 3.3 below. As such, as far as our overall project is concerned, we will adhere to both metasemantical theses of inferentialism.

We will further assume a so-called 'global' version of (metasemantical) inferentialism, that takes not only a specific subset of natural language expressions to be accountable for in inferentialist terms, but for the whole of language. This will naturally entail 'local' inferentialisms with respect to every other subset of expressions, most pertinently the logical vocabulary, if there is one to be found in natural language (ibid., 201), otherwise for the logical vocabulary as it is found in formal logic.[30]

This brings us to the semantical level, where one can distinguish between 'orthodox' and 'unorthodox' semantics. The former kind follows mainstream referentialist

consists in publicly observable linguistic competence with "F".

[30]Why inferentialism about logical vocabulary as it appears in formal logic should be considered as well will be explored in section 4 below. To the extent why this distinction matters, cf. there as well.

truth-conditional semantics and considers inferentialism as a thesis about metasemantics only. The latter sees the rules of inference to be involved on the semantical level as well (cf. ibid., 200). This involvement can take many forms. For example, meanings could be straightforwardly identified with inferential roles or rules of inference, or they could be taken to supervene on the latter, or perhaps such roles and rules merely represent meanings (cf. ibid., 203). As said above, we will take meaning to be at least determined by such roles or rules. However, we can go even further. Recall that in the way I phrased the constitutive argument, it starts with the assumption that meaning is something that can be taught and explained. Yet what we then found to be taught and explained were in fact the rules themselves. As such, we shall take the cue from this and *equate* the meaning of linguistic expressions with the inference rules or roles that govern their use.

It might be contested that this is not necessarily a strong argument. Simply because teaching meaning or knowing meaning is teaching or knowing the relevant rules, does not imply that meaning *is* the relevant set of rules. The counter to this, to my mind, is the question of motivation. If when teaching and knowing the meaning of expressions, one is teaching or knowing certain rules, then what explanatory upshot do truth-values, referents and extensions still provide? Moreover, as other inferentialists have argued before, it seems that truth-conditions face a challenge when taken as a necessary ingredient in meaning, since those truth-conditions could transcend our ability to ascertain the truth-value of the relevant sentence, which seems to square poorly with taking the understanding of a sentence to consist in competence with its use (cf. Miller 2002, 354f.). That is, if we are already committed to the epistemological thesis, it seems we have little room to manoeuvre on the semantical level. Unsurprisingly, this so-called 'manifestation argument' is far from uncontroversial,[31] but it should have made clear that the transition from inferentialist metasemantics to inferentialist semantics is a natural one.

For the remainder of this investigation, then, we will reject mainstream semantics for natural language and assume rules and roles to be *constitutive* of the meanings of expressions themselves. In any case, if we were to only remain on the metasemantical level with inferentialist commitments, yet would adhere to mainstream semantics otherwise, there would (i) be little incentive to explore proof-theoretic conceptions to begin with and (ii) we would be back where we started - namely with the Glanzber-

[31]Cf. Miller (2002) for criticism as well.

gian issues. As such, in order to properly investigate proof-theoretic conceptions, we have to make the step and adopt an unorthodox semantics either way. This will also be in line with the authors which will provide us with the motivation to be inferentialists in the first place, such as Brandom (1983) and Rumfitt (2000).

Moving on, we are put before another decision, namely the one concerning holism. In general, semantic holism about meaning in relation to inferentialism is the thesis that an expression's meaning is "[...] determined by the entire network of inferential connections in which it participates." (Steinberger & Murzi 2017, 201). If such connections are to include even mediate inferential links, it stands to reason that the meaning of one expression is fixed with reference to *all* expressions in a given language. This is undesirable for two reasons. First, our prior commitment to the epistemological thesis above bars us from this option. Holism would entail, on the epistemological level, that in order to acquire one concept I would need to acquire them all (for a given language), given the holistic determination of meaning in this case. Yet this would be quite an incredible feat for somebody not yet acquainted with any language – they would have to learn the language all 'in one go', as it were. Moreover, as Steinberger and Murzi note, this would disqualify ordinary human speakers from understanding any expression in natural language (cf. ibid. 202) – a statement as absurd as it can be. For it is clear that no ascription condition for the ordinary notion of (linguistic) understanding requires us to go this far. Hence, second, this is itself no problem, thankfully, as the entailed epistemological thesis is implausible either way. To use Steinberger's and Murzi's example: "[...] [M]y understanding of 'measurable cardinal' [could not plausibly] be tied to my appreciation of the correctness of the inference from ['Puck is a cat'] to ['Puck is an animal']." (ibid.).

Since semantic atomism – the idea that no possession of a concept requires the prior possession of others – is ruled out by *global* inferentialism by default, given the characterisation of meaning in terms of inferential links, which will inevitably involve more than one concept, we arrive at a middle-ground position. This position holds that we have clusters of concepts that must be mastered *en bloc*, yet do not require the appreciation of all concepts of a language. This position hails from Dummett and is aptly called "molecularism" (cf. ibid., 201). An easy example would be geographical notions such as west, east, south and north, which have to be mastered *en bloc*, but arguably do not *require* conceptual resources from other topics, such as mathematics or taxation laws.

We now arrive at the only decision that we will not have to take a stance on. Inferentialism can be understood either descriptively or normatively. In the former case, the set of inferences that determine the meaning of an expression is given by the inferences speakers actually make or are disposed to make, while in the latter conception the set contains those inferences that we *ought* to draw. Specifically, mastering an expression does not just involve drawing the inferences that the community is drawing, it also consists in recognising the propriety of those inferences, and that propriety is part of the meaning itself (ibid., 202). As far as our concern with logical inferentialism is concerned, I will sidestep the issue of the normativity of logic as far as possible.

This concludes our brief overview of inferentialist themes. As I tried to motivate, we will accept inferentialism as encompassing both metasemantical theses as well as semantical ones. On the metasemantical level, we will hold that competence with inferential rules of use constitute knowledge of the meaning of an expression, and that those rules determine an expression's meaning. Furthermore, we reject truth-conditional and referentialist semantics in favour of non-standard semantics that involve inferential connections and rules, where we reject holism and embrace molecularism. Coming from the constitutive argument, we took the step and identified the meaning of an expression with the inferential rules or roles that govern its use. Since we advocate these theses for natural languages, which are public, we are moreover anti-individualistic. Last but not least, we see conceptual content as at least going hand-in-hand with mastery of public linguistic meaning, and leave the question regarding the normativity of the involved inference rules open.

Before proceeding to the topic of logical inferentialism, which will then lead us over into the next chapter of proof-theoretic conceptions, it is an opportune moment to give proper motivation for being an inferentialist in the first place.

3.3 Why Be an Inferentialist?

Recall that I take the issue of referentialist theories of meaning settled as far as motivation is concerned. The three arguments we saw above seem to lead to the conclusion that both as a primary model of meaning and as a metasemantical thesis about language acquisition, the referentialist picture fails. This does not immediately disqualify truth-conditional semantics though, as noted before. After all, it is possible to give a theory of meaning where truth has a crucial role in the elucidation and constitution of linguistic meaning, despite the absence of referentialist assumptions, as is the case with Davidson (1967).

Robert Brandom, in his paper *Asserting* (1983), allows us to extract an argument that *both* weakens the case for truth-conditional semantics *and* simultaneously strengthens the cause of inferentialism.[32] The general structure of the argument is as follows:

(P1) Assertion is the fundamental activity in which linguistic meaningfulness is manifested.

(P2) A close inspection of our practice of assertion reveals that assertion does not fundamentally require the notions of truth or truth-conditions in order to be elucidated, but can in fact be characterised in purely inferential terms alone.

(C1) (From P1 and P2:) Hence, the fundamental activity in which linguistic meaning is manifested does not require characterisation involving the notions of truth or truth-conditions, but can be characterised purely inferentially.

(P3) If truth or truth-conditions were to play an integral part in linguistic meaning, such notions would be expected to play a crucial role in characterising the practice of assertion.

(C2) (From C1 and P3:) Thus, truth and truth-conditions are superfluous for the elucidation of linguistic meaning, the latter which allows characterisation in inferential terms as far as its manifestation in assertions is concerned.

In other words, and by way of summary, if the fundamental meaning-manifesting activity of assertion can be characterised without the notions of truth or truth-conditions, then what exactly ought to be the explanatory upshot of these notions as far as linguistic meaning is concerned? This is, of course, still somewhat 'gappy' as far as arguments go, but at this point the goal is to merely explicate the main upshot of the argument.

Now, not all of these premisses are explicitly endorsed by Brandom himself. His main concern in the aforementioned paper is not an argumentation against truth-conditional theories, but rather an articulation of his inferentialism. Nevertheless, the premisses not explicitly endorsed or defended by Brandom allow for justification

[32] For this reason, I have opted to discuss Brandom's paper instead of the manifestation argument.

on independent grounds, and it is those to which I will turn first.

(P1) is an assumption Brandom explicitly endorses, but does not justify (cf. Brandom 1983, 637). Is it in need of justification? One might object to the idea that assertion is *the* fundamental activity in which linguistic meaning is manifested. It could be argued that other speech acts, such as questions, or more plausibly, prescriptive utterances are more frequent, hence are more deserving of the label of "fundamental activity". However, this does not establish that assertion is not *a* fundamental activity in which meaning *is* manifested – a claim that is hardly contestable. Moreover, truth-conditional semantics are primarily geared towards explaining linguistic meaning as it appears in declarative sentences, and hence is beholden to play a role in understanding assertion. Hence, even if (P1) is rejected the way it stands, the argument can be easily reformulated in terms of treating assertion as *one* of the fundamental activities, particularly one in which truth-conditional semantics ought to play a role in, anyway. Thus, I hold that (P1) is either acceptable as it stands or can easily be modified to serve the purpose at hand.

This leads us to (P3). Why should (P3) hold? Brandom himself does not comment on (P3) as such, but it can nevertheless be motivated in general. After all, a theory of linguistic meaning ought to supply us, *inter alia*, with a theory of the *content* of assertions. In the context of truth-conditional semantics, those contents are given it terms of truth-conditions, hence we expect assertions to involve those in *some* capacity. For example, we could expect truth-conditions to play a role in deciding which assertions are warranted under what circumstances, or we might expect truth-conditions to be involved in the explanation of what inferences are licensed given certain other assertions. After all, the usual gloss is that asserting that p is putting p forward *as true* (cf. ibid., 638). However, these expectations do not yet go far enough. If truth-conditions are supposed to be an *essential* ingredient to meaning, a claim to which the truth-conditional semanticist seems to be beholden to, truth-conditions *must* play *some* role in the practice of assertion. If they turn out not to, the explanatory ambitions of truth-conditional semantics would be undermined.

Let us now turn to (P2). What reasons do we have to assume that our practice of assertion – including the content of said assertions – can be characterised without using truth-conditions? After all, as Brandom himself agrees, asserting that p is putting p forward as *true* (cf. above). However, he does not stop there, and proceeds by giving an account of what it means to put a sentence forward as true (ibid., 638). The key move, to my mind, is that putting a sentence forward as true is understood as "[...] putting [it] forward as one from which it is appropriate to make inferences. That is, asserting is issuing an inference license." (ibid., 639). As such, despite the customary gloss involving the expression "true", Brandom holds that all that really counts in this regard is the issuing of an inference license. In fact, as he later claims, a speaker taking a sentence as true in the context of the practice of assertion is simply them being prepared to (re-)assert the claim and draw the appropriate conclusions themselves (cf. ibid., 645). In conclusion, the common-sense gloss about assertion involving the notion of truth is scrutinised with respect to the actual doings of speakers, where the notion itself turns out not to play the role a truth-conditional theorist would take it to, but involves the preparedness or appropriateness of drawing certain *inferences*.

This seems to be clearly on the right track as far as the *actual* practice of assertion is concerned. The game of assertion, as a social practice, involves the activities of speakers, which in this case seem to involve the drawing of inferences, not so much an occupation with truth-conditions, much less truth itself.

In order to provide an illustrative example, imagine two people – Alan and Brenda – working in an office. Brenda asks Alan if he has already finished the financial reports of last year, to which Alan responds with "Yes, I have.". The content of Alan's assertion here is that he finished the financial reports of last year, and assuming the orthodox gloss, what Alan did was put the corresponding sentence forward as true. What does this amount to? For the moment, assume that Brenda has no reason to suppose that Alan is lying, or otherwise deceiving her, i.e. Brenda also *takes* the utterance to be true. What could happen? For example, Brenda could tell her boss that Alan finished the reports (reassertion) or she could tell her co-workers not to bother with them, as they have already been done (drawing inferences). She might assert: "You don't need to bother with them, Alan already did it.". This would be an inference drawn from the original assertion, and would follow more or less directly. For what else does it mean that somebody finished their work other than, among other things, that it does not need doing anymore, hence other people need not bother any further? What we see here is that the meaning of a certain phrase, such as "finishing one's work", can be elucidated by studying what role it plays in assertions that contain it, specifically by studying the appropriate inferences that can be drawn from such assertions. This is, in a nutshell, the inferentialist's top-down approach to explanation mentioned in section 3.2. Moving on, what this little example hopefully succeeded in motivating is the idea that as far as the practice of assertion is concerned, putting things forward as true and taking them as true amounts to doing nothing more than licensing further 'assertive work', most prominently drawing inferences.

In order to fully defend (P2), however, it is not enough to merely show that the gloss involving the notion of truth is analysable differently after all. What we need is a full and coherent picture of the practice of assertion, characterised in inferential terms. I will now proceed to give the key points of said picture as far as Brandom expounds them in Brandom (1983). The goal will be to demonstrate that our practice of assertion *can* be characterised in inferential terms, without reference to truth-conditions, and, moreover, coherently so. This will conclude the defence of the premisses of our argument, and hence provide us with the sought-after motivation for (global) inferentialism.

The key points, then, are as follows. First and foremost, if asserting is issuing an inference license, yet inferring is the drawing of conclusions, an assertion amounts to a license for further assertions, namely those inferentially appropriate given the original one. This calls for an account of the appropriateness of inferences, which Brandom takes to be constituted by social norms, i.e. an inference is appropriate to the extent that it conforms with extant social inferential practices of the linguistic community (ibid., 640). To the characterisation of assertion as licensing further assertions – what Brandom labels "endorsement" – the second crucial dimension is that of 'commitment'. If that second dimension were missing, an explication of assertions in terms of inferences would arguably fall short, in virtue of taking "[...] us around in a rather small circle." (ibid., 641). As Brandom continues, the practice of assertion not only involves the licensing of further assertions taking the original one as a premise, but also involves a commitment of the speaker to justify or vindicate the original one. If such a defence fails, the original assertion loses its social significance (cf. ibid., 641). Of course, this defence is only needed if the original claim is challenged. Concerning the defence of a claim itself, it is clear that such a justification will proceed by producing further assertions, namely those that taken as premisses allow one to appropriately infer the original claim (ibid., 642). Thus, inference is the key notion that underpins our practice of assertion, specifically those inferences deemed appropriate by the linguistic community. As such, truth-conditions and truth itself seem to fall out of the picture at this point.[33]

We have already seen an example of endorsement in our office-example above. With his assertion that he has done the work, Alan endorsed this claim. He made it possible for Brenda to reassert it, or advice people on the basis of it, by drawing appropriate (practical) inferences from it. But Alan is also committed to this assertion. If Brenda were to retort to the original claim with "Are you sure? I only saw you browsing the internet this morning.", he is expected to defend his original claim, for example by claiming that he did the work some other time, or by claiming that Brenda just happened to look over Alan's shoulder when he was surfing online, missing the times he put in actual work. If Alan would turn out to have no good counterargument, Brenda would not be disposed to further reassert Alan's claim, nor would she continue to advise her co-workers on the basis of it. In other words, Alan's failure to defend his claim would rob it of its authority. Of course, the most obvious defence of Alan's claim is for Alan to simply demonstrate that the work is done, for example by showing the finished reports to Brenda. As such, we must be reminded that *sometimes*, what counts as grounds for assertion of a sentence might be a verification, and in this sense 'truth-conditions' *can* play a role. Nevertheless, they form a special case as grounds for a warranted assertion, and do not play a general role in determining meaning.

These two pillars – endorsement and commitment – are further expanded upon by recognising the social institution of deferring one's responsibility of justification of a claim to the speaker that originally made it (ibid.). Thus, assertion as a practice takes place in a web of authority and responsibilities (ibid., 642f.). Once more pertinent to our goal of getting a motivation for the inferentialist project as a whole, Brandom remarks that by being caught-up in this web, sentences acquire content in terms of an inferential-justificatory role, a role which in order to be grasped requires an understanding of what counts as a justification for the sentence, and what can be correctly inferred from it, in accordance with the accepted standards of the linguistic community (cf. also Brandom 2000, 63). What is crucial here is the claim that failure to grasp these aspects of its inferential-justificatory role precludes a speaker from asserting the sentence in the first place (Brandom 1983, 643). As such, once again, it is not knowledge of its truth-conditions that enable an assertion of a sentence, but

[33]This, of course, does not disqualify truth as a standard for rational discourse. The issue is entirely with truth as an essential ingredient in meaning – in the form of truth-conditions, specifically.

knowledge of its inferential-justificatory role that does so. Of course, the immediate counter of the truth-conditional theorist is to claim that this role is obtained via considerations of truth-conditions. We will return to this point shortly. Brandom himself, in any case, takes the norms of the linguistic community to be the only authority in this regard (cf. ibid., 644).

Both deferrals and how sentences obtain a justificatory-inferential role have already been showcased in our accompanying example. When Brenda's reassertion of Alan's claim would be challenged by her boss, the straightforward reaction for Brenda would be to simply defer to Alan – "Well, Alan said so.". The more interesting aspect is arguably the idea that sentences gain content in this web of inferential practices, so let us look at the sentence "I [i.e. Alan] have finished the financial reports of last year.". The meaning of this sentence, so Brandom claims, is determined by what constitutes an appropriate justification for it, and by what counts as appropriate inferences from it. We can see how this works in this case, by looking at how the content of an assertion of it is related to both of these aspects. Assume Alan insists that his having spent the entire morning browsing the internet counts as a justification for his utterance "I have finished the financial reports of last year.", or that he keeps telling his co-workers that they must finish those reports still. If we – i.e. the linguistic community – were to be confronted with this scenario, would we still be willing to say that Alan's purported assertion has content? If it does, then clearly it does not have the content that the community at large would impute to it. In fact, we would be prepared to say that what Alan is saying is not that he finished the financial reports, but actually the opposite, i.e. that the work is *un*finished. Moreover, if we wished to correct Alan on this account, we would arguably proceed by telling him what people usually take this sentence to mean, by means of explaining what counts as grounds for a warranted assertion of it and what can be appropriately inferred from it. As such, on the level of assertions, their content does indeed seem to be determined by what the linguistic community at large treats as appropriate justifications for it and appropriate consequences from it. In any case, the fact that Alan would associate those inferences with the sentence would rob his assertions of it of social significance, as we could not impute to Alan that he knows what he is claiming. This last point also demonstrates once more how knowledge of meaning is knowledge of correct use, this time in the more detailed form of justificatory-inferential roles.

As a last ingredient to the game of assertion, we need an account of what it means to take somebody else's claim to be justified. As Brandom puts it:

> [Taking somebody else's claim to be justified] is for the respondent to recognise the inferential authority of the original remark. This recognition in turn consists in the respondent's disposition to accept as legitimate deferrals of justificatory responsibility to the original assertor.
> (ibid., 645)

In other words, to take an assertor's claim to be justified is to be ready to defer the responsibility of justifying the claim to the assertor in case that claim is challenged. In order to close off our example, we already saw that unless Brenda had good reasons to suppose otherwise, for example by seeing Alan procrastinate instead of work, she will be ready to defer to Alan if anybody were to challenge the claim. If she were not ready to do so, we would clearly not say that Brenda takes Alan's claim to be justified, while we would arguably assume Brenda to take it so if we see her deferring to Alan. In other words, we would see Brenda as taking Alan's claim to be justified just in case she exhibits the relevant inferential behaviour.

These four moves then constitute the game of assertion, according to Brandom. They consist in *making an assertion* in the first place, i.e. committing oneself to the statement and endorsing it, *challenging* the assertions of others, asking for further assertions that support the original claim, *defer* the justificatory responsibility to other speakers, and *recognise* other claims as justified. *All* these moves were explicated without reference to truth-conditions, and only with reference to the nature of social practices, standards and the correctness of inferences, the latter which were explicated in terms of the former two. We can now address the earlier rebuttal of the truth-conditional theorist: What explanatory upshot do truth-conditions provide, for example by claiming that they implicitly determine the inferential-justificatory role of a sentence after all, if all that is manifested in the actual practice is conformity to the linguistic community's standards? It seems, then, that truth-conditions are superfluous, and inferences, whose propriety is determined by the social norms of the linguistic community alone, are the notion we really need in order to describe the practice of assertion – as well as its contents, which are determined by their role in the former.

This, then, concludes the argument against truth-conditional semantics that simultaneously ought to motivate inferentialism.[34] It should come as no surprise that Brandom's variant of inferentialism is not without its critics.[35] However, as I said in the beginning of this chapter, it was not the goal to settle the debate surrounding inferentialism, but to simply provide a motivation for it, as well as reasons to doubt the mainstream referentialist truth-conditional approaches. It hopefully succeeds in motivating the explanatory metasemantical thesis as well as that rules of inference and justificatory roles are, in some sense, constitutive of meaning. Given our earlier reflections on the matter, we shall take this as sufficient motivation to further equate meaning with the pertinent rules.

Still, an unsatisfied reader might maintain that these motivations so far are rather general, and she would be right in pointing this out. For what we are after in the end are proof-theoretic conceptions of *logical* notions, and little has been said so far on the inferentialist's take on purportedly logical vocabulary in natural language, or clearly logical vocabulary within the domain of formal logic. Thus, we close this chapter by considering logical inferentialism, while keeping in mind, however, that for the ultimate goal of a comparison between natural language and proof-theoretic conceptions, the previous lessons will be just as paramount.

3.4 Logical Inferentialism and Its Motivation

3.4.1 The Motivation: Underspecified Senses and Logical Competence

I wish to reverse the order in this section, and start with the motivation for logical inferentialism, before expounding its commitments in detail. The reason for this is twofold. On the one hand, most readers will already be acquainted with the standard

[34] As Steinberger and Murzi (2017, 211) point out, this picture shows that Brandom maintains a normative understanding of inferentialism. Since we are using his account of assertion as a motivation to be inferentialists, yet I claimed to try and sidestep the issue of the normativity of logic, this raises the question whether this is still possible. As Brandom remarks, the normativity of his account stems from the general normativity inherent in social practices as manifested by authorities, responsibilities and obligations (cf. Brandom 1987, 640). This kind of normativity clearly transfers onto the use of (purportedly) logical vocabulary in natural language, but is tied to *adhering to the rules* of use and recognising their authority *as* rules of use, whereas the issue of the normativity of logic knows many variants and worries (cf. Russell 2020, sect. 4). As such, this kind of 'semantic normativity' that is inherent in Brandom's view might square well with views that think that logical laws are devoid of normativity, as the semantic normativity in question has nothing special to do with logic, but with adherence to rules in general. Thus, I claim, we can sidestep this issue for the remainder of this investigation.

[35] Cf. MacFarlane (2010), for a pertinent example.

picture of the meaning of sentential connectives, namely as truth-functions. For example, the connective "and" gets the following treatment:

$$(\&) \qquad [\![\text{“and”}]\!] = f(t_1, t_2) := \begin{cases} T, & t_1 = T \text{ and } t_2 = T \\ F, & \text{otherwise} \end{cases}$$

i.e. "and" takes two truth-values and maps them to the value 'True' just in case both inputs were 'True', and to 'False' otherwise. On the other hand, this time the motivation and subsequent formulation of inferentialism go hand-in-hand, and as such allow for a natural transition from the topic of motivation to the content of the position.

Ian Rumfitt (2000), gives us the following argument to be an inferentialist about logical vocabulary. He starts by asking what the (Fregean) sense of a sentential connective consists in, i.e. "[...] the logically relevant part of [its] meaning" (Rumfitt 2000, 782). By way of illustration and elucidation, he considers the classical truth-functional approach as it is exemplified by (&) above. He agrees that such clauses succeed in capturing the relevant logical meaning of these connectives, at least in a sense (ibid.). What such clauses do not address, however, is what makes such clauses interpretative for speakers of, in this case, English.[36] Additionally, he contrasts this with other connectives such as "but", where it is usually agreed upon that, while "but" and "and" do not have the same meaning, they are nevertheless alike when it comes to the *logically relevant* parts of their meaning, i.e. their senses, as Rumfitt understands the term. As such, he wishes to inquire in what *this kind* of sense consists in (ibid., 783).

It is here that Rumfitt makes an argument against the kind of standard account I have indicated above with (&). He takes this standard account to claim that the logically relevant part of the meaning of a connective is captured and specified by supplying the relevant truth-table. Moreover, a speaker is said to understand or grasp the connective's sense if and only if they have knowledge of the corresponding truth-table (ibid., 784). Thus, for example, a speaker grasps the sense of "if and only if" just in case they know the following truth-table:

[36]In this vein, recall Hanfling's observations from chapter 1.

A	B	A if and only if B
T	T	T
T	F	F
F	T	F
F	F	T

In addition to this, this truth-table fully determines the (logically relevant) sense of the connective. However, as Rumfitt points out, consider now the following truth table for the 'new' connective "conn":

A	B	A conn B
T	T	T
T	F	T
F	T	T
F	F	F

According to the standard account, the relevant logical sense of "conn" should therefore have been determined, and a speaker grasps that sense just in case they have knowledge of the above truth-table. The issue now, though, is that this truth-table alone does not seem to tell us whether "A conn B" means the same as "A or B", with "or" in its inclusive sense, or the same as "It is not the case that A is not the case and B is not the case", or any other logically equivalent sentence to "A conn B". In other words, grasping this truth-table does not seem to allow a speaker to distinguish between all these different sentences, nor does the truth-table itself seem to distinguish between them, even though they intuitively appear to differ in some logically interesting aspects (cf. 784f.).

Then again, it might be contested that this is precisely what has been argued for. It might be conceded that there are minute differences in meaning between, for example, "A or B" and "It is not the case that A is not the case and B is not the case" even with respect to their logical parts, for example because one truth-table takes some extra work to construct, but that they nonetheless have the same sense, since they ultimately share the same truth-table. However, corresponding inferences involving them are clearly of a different nature. Consider an inference of the form "Two is prime, hence two is prime or four is prime" and "Two is prime, hence it is not the case that two is not prime and that four is not prime". The first inference seems to be a seamless one: There is no way of further breaking it apart into (even) smaller inferential steps. This is quite different in case of the latter inference, which can be broken apart into an application of *reductio ad absurdum* (cf. ibid., 785). As Rumfitt claims:

> The fact that [an argument] proceeds by a series of simple inferential steps is surely a feature that ought to interest the logician. But it is a feature that may be disturbed by substituting an instance of ["It is not the case that A is not the case and B is not the case"] for the corresponding instance of ["A or B"].
> (ibid.)

Hence, there seems to be a logically interesting difference between sentences like "A conn B", "A or B" and "It is not the case that A is not the case and B is not the case", which is given by how an argument might change from 'gap-free' to 'gappy' when one sentence is substituted for the other, *despite* them sharing a truth-table. Thus, it seems that truth-tables are not capturing all logically interesting aspects of a connective's meaning (ibid., 785f.).

As a last line of defence, it might be suggested that the standard account could be amended to treat only 'semantically primitive' connectives as having their sense determined by a truth-table. After all, the semantic structure of "A and B" and "It is not the case that, A is not the case or B is not the case" clearly differ in their complexity. Hence, "[...] [t]he only cases which would threaten [this account] are those in which two unstructured connectives share a truth-table while differing in sense." (ibid., 786). However, such examples are easy to come by, as an example due to Peacocke shows (cf. Peacocke 1987, 156). Assume that the unstructured connective "$\otimes$" is such that anybody who understands it finds it 'primitively obvious'[37] that "A $\otimes$ B" is incompatible with both "It is not the case that A" and "It is not the case that B". It is clear then that, assuming classicality, "A $\otimes$ B" is equivalent to "A and B", hence they share a truth-table. However, it is clear that the inferential steps involved in verifying said equivalence are more than one, hence these two differ in sense after all (cf. Rumfitt 2000, 786).

In summary, it seems as if the truth-table based approach is simply expressively

[37] I will return to this notion shortly. For the time being, I am relying on the reader to have an intuitive idea of it.

poorer than the corresponding *inferential* approach, suggesting that we should forgo the former in favour of the latter. We obtain a more expressive tool-kit for the logician's work when adapting an inferentially-based account, which in turn allows us to capture aspects of a connective's meaning that do seem to be of interest to the logician, which would be absent in a truth-table-based approach.

As a precursor to chapter 4, Rumfitt makes the following additional remark. He conjectures that the fact the philosophers of logic have missed this important difference is because of the prevalent notion of validity, namely the one that treats an argument valid just in case it preserves truth – or whatever value is the designated one. This is, of course, our old companion (III). If furthermore logic is taken to be concerned only with validity, then whether an argument includes 'jumps' or proceeds by 'gap-free' inference steps is irrelevant. Given the above considerations, it is thus unsurprising that a connective's sense could be thought to be specified by a truth-table alone. Rumfitt objects to this notion of validity with the following example of an argument scheme, due to Shoesmith and Smiley (1978, 105):

(P1) All a are b. Assumption.

(P2) All b are c. Assumption.

(C1) All c are b. Inferred fallaciously from (P2).

(C2) All a are c. Inferred fallaciously from (P1) and (C1).

Both inferential steps are classic fallacies in quantificational logic, and yet, according to the received notion of validity, (C2) does follow from (P1), (P2) and (C1), since there is no model in which all three statements would be true yet (C2) would fail to be so. As Shoesmith and Smiley further argue, this cannot be rectified by requiring all simple inferential steps to be valid individually, as this alone would validate the following disastrous scheme (cf. ibid.):

(P1) A Assumption.

(P2) $A \rightarrow B$ Inferred validly from (P3).

(P3) $A \wedge (A \rightarrow B)$ Inferred validly from (P1) and (P2).

(C) B Inferred validly from (P3).

As such, not only do the individual inferential steps need to be valid, they must also be in the correct order, i.e. we should be able to exclude argumentative circles as in the last scheme (cf. ibid.). This is, *inter alia*, a direct motivation to adopt the proof-theoretic conception of logical consequence, to which we will return in due time. It is important to be showcased now already, though, as it arguably also further motivates logical inferentialism, given the connection between the treatment of connectives and the received view of validity.

A further argument in favour of logical inferentialism is the following, which relies on ascription conditions for competence with logical vocabulary (cf. Steinberger & Murzi 2017, 205). For example, it might be that somebody who fails to correctly apply the following inference rules could not be credited with understanding the word "and" (cf. also Boghossian 2011, 493):

$$\frac{\text{A} \qquad \text{B}}{\text{A and B}} \qquad\qquad \frac{\text{A and B}}{\text{A}} \qquad\qquad \frac{\text{A and B}}{\text{B}}$$

where the horizontal stroke is to be read as "hence". Meanwhile, somebody who has mastered these rules could not reasonably be denied to have understood the word "and". Once again, we see a motivation to adopt an inferentialist stance on the meaning of a set of expressions due to epistemological considerations.

In the end, the goal was once more to motivate, not to vindicate. I was hopefully able to showcase such motivation, and thus it is time to delve into the details of logical inferentialism itself.

3.4.2 Logical Inferentialism Proper: Inference Rules and Harmony

After that has been said so far, the reader should already have a good idea of what logical inferentialism ultimately amounts to. In a nutshell, it is the view that takes the meaning of logical vocabulary to be given by the rules of inference that govern their use in deductive reasoning. For example, Rumfitt says the following:

> Under this [approach], one gives a connective's sense – not by directly assigning any semantical value to it – but by characterising a putatively basic aspect of its deductive *use*.
> (Rumfitt 2000, 788, emphasis in the original)

In the same spirit, Steinberger and Murzi claim the following:

> Logical inferentialism [...] becomes the claim that the meaning of logical expressions are fully determined by their [inference rules], and that to understand such expressions is to use them accordingly to such rules [...]. (Steinberger & Murzi 2017, 204)

We have already seen examples above. One could take the meaning of "if"-"then" to be rules such as *conditional proof* and *modus ponens*, or the meaning of "and" to be the standard rules of conjunction introduction and elimination.[38] Since we made the decision to treat meaning as being constituted by (or to be identified with) rules or roles of inference, we will take logical inferentialism to be the view that the meaning of logical vocabulary *are* these kinds of rules.

Concerning introduction and elimination rules, a reoccurring theme in inferentialism is the distinction between the conditions under which a sentence might be correctly asserted, and what can be correctly inferred from an assertion of a sentence (cf. ibid., and the discussion of Brandom above). As Brandom argues, if we give primacy to the grounds of assertion, a trained parrot who reliably utters "This is red!" when presented with red things would have to be attributed a grasp of the concept 'red'. However, the reason we do not do so is precisely because the parrot cannot infer from its report that this thing is coloured, or connect it with other pertinent notions via appropriate inferences. If were to assign primacy to the appropriate consequences of an assertion, on the other hand, I might end up knowing all the appropriate consequences to be drawn from an assertion of "This is morally wrong." without ever being able to make such an assertion myself in the first place. Yet possession of such moral concepts arguably require a speaker to also know when an act is morally wrong to begin with (Brandom 2000, 64ff.).

This theme finds its most natural expression in natural deduction systems, where each connective is specified with its associated introduction and elimination rules. The idea is to read the introduction rules as specifying the conditions under which a sentence of the relevant form can be introduced within a proof, while the elimination rules specify what consequences can be drawn from the complex sentence. We have already seen such rules spelled out above in the case of "and". If we read "$\rightarrow$" as standing for "If..., then...", we could moreover give the rules for this pair – conditional proof (CP) and *modus ponens* (MP) – as follows:

[38] In the latter case, we will return to this option in detail in chapter 5.

$$\dfrac{\begin{matrix}[\mathrm{A}]\\ \mathcal{D}\\ \mathrm{B}\end{matrix}}{\mathrm{A} \rightarrow \mathrm{B}}\ \mathrm{CP} \qquad\qquad \dfrac{\begin{matrix}\mathcal{D}_1\\ \mathrm{A}\end{matrix} \qquad \begin{matrix}\mathcal{D}_2\\ \mathrm{A} \rightarrow \mathrm{B}\end{matrix}}{\mathrm{B}}\ \mathrm{MP}$$

where $\mathcal{D}$, $\mathcal{D}_1$ and $\mathcal{D}_2$ stand for some further deductive reasoning. Many other authors have made the same assumption in taking the relevant inference rules to be those of natural deduction systems (cf. Rumfitt 2000, 788, and Dummett 1991, 185). As such, it is fair to say that the standard way of specifying the inference rules that constitute a connective's meaning are its standard introduction and elimination rules of the salient natural deduction calculus.

However, this raises a simple question: What makes these rules so suitable, but not others? I have already remarked that the general idea of using introduction and elimination rules flows from the theme of assertion conditions and their consequences. However, this alone does not tell us which rules we ought to take as introduction and elimination rules, apart from the simple requirement that the sentence derived at the end of an introduction rule ought to feature the salient connective as the principal one, and that the elimination rule ought to eliminate that sentence which features the connective in principal position, accordingly. To this end, Rumfitt uses a notion due to Peacocke (1987), namely the previously encountered notion of 'primitive obviousness'. Rumfitt defines it as follows:

> [A rule of inference] is primitively obvious if[f] (a) its soundness is obvious to anybody who understands the [connective(s)] that figure(s) in it; and [...] (b) that obviousness 'is primitive in the sense that it is not consequential upon his acceptance of some more primitive principle; nor upon iterated application of any single principle; nor upon any belief not already presupposed in grasp of the component contents' – i.e. not already presupposed in understanding the various particular instances of the rule.
> (Rumfitt 2000, 787, cf. also Peacocke 1987, 154)

If we take this (lengthy) criterion to heart, we can see that in the case of "if"-"then"-sentences, *modus ponens* arguably fits the description of primitive obviousness. In fact, it is plausible that in the case of conditionals, competence with *modus ponens* is part and parcel of the ascription conditions for competence with such sentences, hence, due to conceptual reasons, the rule could not fail to be obvious to anybody already established to be competent with the connective. Thus, we can see why the rules of natural deduction systems make excellent candidates: They are already specifically chosen to be as primitive as possible from a deductive point of view, satisfying (b) as a result. Moreover, (a) is, in a sense, 'trivially' satisfied assuming that the proof theorist is a competent speaker of their language, and models their inference rules on the basis of it. Lastly, primitive obviousness also guarantees gap-free inference rules, in virtue of condition (b), as 'gappyness' would be revealed by deriving the purported rule by using other, more basic ones, hence violating at least one of three conditions immediately.

What Rumfitt does not explicate is the notion of obviousness itself, however. We can explicate this notion in an intuitive way by seeing it as a matter of a disposition to accept the inferences when presented with them. In this sense, an agent treats a rule of inference as obvious just in case they are disposed to assent, affirm or otherwise accept the inference rule whenever instantiated, without the need for justification or elaboration. Arguably, in most cases this will be manifested in simple non-objecting behaviour on part of the agent. Just as Brandom reminds us, in ordinary discourse, claims – and the inferences linked to them – are treated as appropriate or justified *unless* explicitly challenged (cf. Brandom 1983, 642).

Let us summarise our findings so far. Logical inferentialism is the position that treats the meaning of logical vocabulary to be constituted by the rules of inference that govern their deductive use. Specifically those rules are chosen that are primitively obvious, which, as we saw, are generally corresponding to those of natural deduction systems' introduction and elimination rules – rules which accord well with inferentialism's distinction of assertion conditions for a given sentence and the appropriate consequences that can be drawn from it. Moreover, in virtue of being primitively obvious, these rules are gap-free, in the sense of it not being possible to break the relevant inference apart into smaller, more basic steps.

Before moving on, I wish to address a worry that the attentive reader might have noticed. In the examples so far, I have tended to shift between natural language inference patterns and the more formal ones of formal logic. By doing this, I have in part simply followed the authors I have been discussing. However, the obvious worry is that by not properly distinguishing between natural language and formal logic, I might ultimately be begging the question in relation to our ultimate goal, namely determining the truth of the logic in natural language thesis (LNLT) in the case of proof theory and inferentialism. As such, I wish to remark at this point that while I think that the danger is real, nothing said so far has been an error in this regard. First, the example of "and' has been a deliberate choice in virtue of being less contentious than, for example, "or" and disjunction's standard introduction rule: $A \vdash A \vee B$, as other authors would be happy to point out (cf., for example, Strawson 1952, 90ff.). Second, and more importantly, we will take a step back in chapter 5 and get clear about proof-theoretic commitments in relation to constants and consequence relations, *before* approaching natural language *again* without prior assumptions about the relation between proof theory and natural language. In the end, all that inferentialism about natural language commits us to is to treat inferences as the primary explanatory notion of meaning. Whether we will find the kind of inferences centred around certain words in natural language that properly classify as *logical* is still left open, and to be determined with reference to proof-theoretic conceptions. In addition to this, logical inferentialism can easily be seen as a position about the meaning of logical vocabulary as it appears in *formal logic*, though naturally in close connection to natural language reasoning, where logical inferentialism's motivation comes from. Still, at this point it is entirely possible that inferentialism is mistaken in taking introduction and elimination rules as they appear in formal systems to determine the meanings of *natural language* expressions such as "or" or "if and only if", even if they manage to properly capture the meaning of "$\vee$" and "$\leftrightarrow$".

Moving forward, the message to be reminded of here is that logical inferentialism might already incur some commitment related to the LNLT, and cannot be taken for granted to the extent that it does. With respect to said extent, we can moreover treat logical inferentialism as a thesis about the meaning of the connectives as their appear in formal logic first and foremost, and suspend judgement about the relation to natural language for the time being. It is, of course, the close connection between the content of logical inferentialism and its motivation stemming from more ordinary, i.e. not already formalised, forms of reasoning that make the LNLT more plausible in our case to begin with.

Let us now, before moving on to proof theory proper, discuss one last aspect of logical inferentialism. No discussion of logical inferentialism would be complete without it, and so we must turn to the notion of harmony.

As Steinberger and Murzi point out, logical inferentialism is a kind of conventionalism, in the sense that logical laws are ultimately taken to be founded in our practices, and are hence thought to hold ‘by convention’, rather than by answering to any realm of abstract facts (cf. Steinberger & Murzi 2017, 206). Quoting Dummett, they explicate the general commitment of the inferentialist as maintaining that logical laws can be justified only to the degree that they can be reduced to basic laws, which in turn, on pain of regress, cannot enjoy the same kind of justification. This, now, raises the question of whether any purported basic law can be seen as determinative of meaning. As is well-known, any affirmative answer to this question yields catastrophic consequences.

Arthur Prior (1960) proposes the connective “tonk”, whose rules are a combination of the introduction rule for disjunction and the elimination rule for conjunction:

$$\frac{\text{A}}{\text{A tonk B}}\ \text{tonk-I} \qquad\qquad \frac{\text{A tonk B}}{\text{B}}\ \text{tonk-E}$$

Adding “tonk” to any derivation system allows one derive any theorem, provided one can prove at least one, thanks to the possibility of chaining the I- and E-rule.

The standard solution to this problem is the notion of *harmony*, anticipated by Gentzen (1934) and made popular by Dummett (1991) (cf. Steinberger & Murzi 2017, 206). The initial idea of Gentzen was to see the I-rules as laying down the definitions of a connective, by specifying the conditions under which a sentence containing the connective in principal position may be asserted, and to treat the E-rules as consequences of those definitions, specifically in the sense that the elimination rule is beholden to the ‘resources’ provided by the introduction rule (cf. Steinberger & Murzi 2017, 206, and Gentzen 1934, 80). This thought can naturally be reversed, should the elimination rule be seen as more intuitive or constitutive of meaning, as it might be the case with *modus ponens*. As a result, I- (or E-) rules cannot be chosen arbitrarily, once the other one has been fixed (cf. Steinberger & Murzi 2017, 206f., and Dummett 1991, 215). The idea, in a nutshell, is this. There ought to be a kind of equilibrium between the grounds on which we can introduce a complex sentence as according to an introduction rule and the consequences that can be drawn from such an introduction (cf. Steinberger 2009, 62). More precisely, we require that the elimination rules do not license inferences that go beyond what the conditions stated in the introduction rules allow, nor should the elimination rules prevent us from being able to derive consequences that would be licensed by the conditions found in the corresponding introduction rules (cf. ibid., 62f.).

Arguably the standard way of spelling out this intuitive idea is to use peak-levelling and expansion criteria. A *local peak* inside a derivation is taken to be a formula containing a connective c in principal position that has been introduced via its introduction rule, yet is then immediately eliminated again via its elimination rule. The condition now states that such a peak must be able to be levelled, i.e. the derivation must be transformable into one with the same or fewer assumptions and the same conclusion, but without the unnecessary detour over the peak. This ought to rule out the case where elimination rules go beyond their corresponding introduction rules. For if that were the case, levelling a peak while retaining the same conclusion could not work, for the overly permissive elimination rule will generate additional conclusions that will remain after the levelling. However, we also need a safeguard against the opposite vice, where the elimination rules are too restrictive. To this end, we must further require that a derivation that terminates in a formula containing the connective c in principal position can be extended via a detour of, first, elimination and, then, introduction rules into a longer derivation with the same assumptions and conclusions. This is due to the fact that if the elimination rules are too restrictive, they would fail to make full use of the conditions laid out in the I-rules, hence we will be unable to chain the latter to the former. In conclusion, then, the rules for a connective c are harmonious just in case there are both levelling and expansion procedures (Steinberger & Murzi 2017, 207).

It might be instructive to have a clear example for both procedures. To this end, consider the standard implication rules as seen above. As an example of peak-levelling, assume we have a derivation with (up to the point of the peak) open assumptions contained in the sets Γ_0 and Γ_1, which in the former case lead to the application of *conditional proof* (CP) via some derivation $\mathcal{D}_0$. Furthermore, from Γ_1, we obtain A via $\mathcal{D}_1$. The conclusions are then discharged again via the application of *modus ponens* (MP), yielding the conclusion B:

$$\dfrac{\dfrac{\begin{matrix}\Gamma_0, [A]^1 \\ \mathcal{D}_0 \\ B\end{matrix}}{A \to B}\,\mathrm{CP}_1 \qquad \begin{matrix}\Gamma_1 \\ \mathcal{D}_1 \\ A\end{matrix}}{B}\,\mathrm{MP}$$

We can now easily level this peak by substituting the open assumption A in the application for (CP) with the derivation of A from Γ_1, reducing the open assumptions to those in $\Gamma_0 \cup \Gamma_1$, yielding the simple derivation:

$$\begin{array}{ccc} \Gamma_1 & & \\ \mathcal{D}_1 & & \\ A & & \Gamma_0 \\ \hdashline & \mathcal{D}_0 & \\ & B & \end{array}$$

(cf. Steinberger & Murzi 2017, 207). To see the corresponding expansion procedure, consider now a derivation with open assumptions Γ leading to the conclusion $A \to B$:

$$\begin{array}{c} \Gamma \\ \mathcal{D} \\ A \to B \end{array}$$

Such a derivation can now always be expanded into one that has the same conclusion and assumptions, yet involving a detour via an application of both $\to$E and $\to$I:

$$\dfrac{\dfrac{\begin{array}{c} \Gamma \\ \mathcal{D} \\ A \to B \end{array} \qquad [A]^1}{B} \to\mathrm{E}}{A \to B} \to\mathrm{I}_1$$

Having shown that both expansion and peak-levelling procedures are available, we can conclude that the rules of material implication are in harmony with each other.

At this point in the discussion, the question of motivation is once more raised, however. Why *should* introduction and elimination rules be harmonious at all, *apart* from safeguarding us against examples like "tonk"? One central motivation that Steinberger and Murzi give us is the idea that logic is inherently innocent, i.e. "[...] it shouldn't allow us to prove atomic sentences that we couldn't otherwise prove by first introducing and subsequently eliminating a logical operator [...]" (ibid.). In other words, logic alone should not make it possible to make substantive claims about the world or gain new information 'out of nowhere' (cf. Steinberger 2009, 60). If the rules of a connective would fail to be harmonious, this would ultimately allow for such derivations, as the example of "tonk" vividly demonstrates. As such, in this regard the idea would be to treat harmony as a necessary condition for logicality.

However, as Rumfitt (2017) argues, the principle of innocence is questionable. He gives the following counter-example. He considers the case where I know, through astronomical theory or some other resource, that a certain celestial body B is either in region R or in a black hole. Suppose that after further observations I come to conclude that B is not in R. I can then conclude that B is in a black hole, even though it would be impossible for me to verify this, much less discover it in any other (empirical) way (ibid., 239). Hence, Rumfitt concludes, "[...] the principle of innocence is far less compelling than Steinberger supposes." (ibid., 238). In addition to this, even if we were to accept the principle, the weaker requirement of global conservativeness would suffice to secure innocence. As Rumfitt points out, this is satisfied if *all* the logical rules *together* do not create new grounds for asserting non-logical sentences. As such, there is no reason to require further that I- and E-rules match on the local level. Harmony is, in a sense, 'overkill' (cf. ibid., 239). For these reasons, we shall not adopt harmony as a necessary criterion for logical constanthood.[39] We will return to the issue of conservativeness in chapter 4.

After this short exposition of harmony, we have completed our journey through the topic of inferentialism. We have seen the general idea behind inferentialism, and specifically behind logical inferentialism, as theories of meaning concerning either natural languages (and potentially certain sub-parts of it) or formal languages, specifically logical ones. With these general meaning-theoretic observations completed, I will now set out to finally delve into proof-theoretic conceptions of logical notions proper, which will then complete the set-up for ascertaining the status of the LNLT.

[39]It might be thought that harmony might still be a sufficient condition. However, this is also called into question, as under such an approach certain mathematical expressions will turn out to be logical (cf. Steinberger 2009, fn. 10). Clearly, we do not need to make any decision in this regard, nor should we, hence I shall reject harmony as a sufficient criterion as well. This will hopefully become clearer once we discuss the purely inferential character of logical constants' meaning-constitutive rules, for being purely inferential seems to square badly with being mathematical, hence not being entirely topic-neutral. Cf. section 4.2 below.

4 Proof-Theoretic Conceptions of Logical Notions

We now arrive at the discussion of proof-theoretic conceptions of logical consequence and logical constants. This short chapter is organised as follows. I will start with some general remarks about the situation concerning proof-theoretic conceptions and formal inferentialist semantics, and the implications about how to proceed, accordingly. Then, I will be spelling out the idea of the proof-theoretic conception of logical constants, which will be followed up with what it means for a conclusion to follow logically from a set of premisses. As we will see, the later characterisation *depends* on the former. At the end of the chapter, we will finally excavate the realisability conditions, which we will rely on in the final chapter.

4.1 Methodological Considerations

There are three general issues with respect to the content of this chapter, relating to the details of the proof-theoretic conceptions, their general conceptual underpinnings, and the current status of proof-theoretic semantics. These will have a methodological impact on how we should proceed in this chapter and beyond, thus I shall discuss them first.

First, the specific details on how we conceptualise logical constants and consequence relations under the prism of proof theory are largely irrelevant to ascertaining the status of the logic in natural language thesis (LNLT) with respect to inferentialism. This is due to the fact that if we had doubts that the general idea behind proof-theoretic conceptions would square badly with our inferentialist (meta)semantics, then there would arguably be little further incentive to check the details. Of course, this does not imply that the details are *fully* irrelevant. For it is possible that while the general idea matches well with our meaning-theoretic assumptions, the devil might hide in those very details, which might in turn render the LNLT false. However, there are two further reasons as to why the details of the proof-theoretic conceptions are of minor, if not negligible, importance.

The second reason has to do with the lesson we learned at the end of chapter 2. Recall that the conclusion was that contemporary practice in formal semantics seems insufficient to determine the status of the LNLT. Even though the set-up of contemporary referentialist truth-conditional semantics and model-theoretic conceptions allowed for direct comparison, the practice of the former gave no conclusive evidence with regards to the LNLT. This was due to the fact that the only kinds of entities that would seem to motivate a reintroduction of a sufficiently large number of models into the framework of absolute semantics are precisely the *logical* ones, hence the status of the LNLT cannot be determined with reference to contemporary practice in formal semantics alone. Thus, we may doubt that spelling out the details on the formal side of things, while facilitating more direct comparison, might leave us equally clueless in our case of inferentialism and proof-theoretic conceptions, since the status of the LNLT might end up equally independent of the current practice in proof-theoretic semantics.

This is further compounded by the third reason. As it stands, the project of proof-theoretic semantics is still in its infancy, being mostly focussed on accounting for the meaning of logical constants in terms of proofs (cf. Schroeder-Heister 2018, introduction and sect. 4), and most (general) inferentialists seem to stay on the conceptual level, eschewing work in (proof-theoretic) formal semantics. As such, *even* if we *had* a fully fleshed out proof-theoretic conception at hand that accords well with our meaning-theoretic assumptions, the comparison work would leave much to be desired, as large parts of natural language have not yet been accounted for in proof-theoretic semantics.

This last point deserves further emphasis. In contrast to Glanzberg's set-up, our starting point is different. This puts our project in a different light, for we cannot straightforwardly compare inferentialist approaches to linguistic meaning with proof-theoretic conceptions, since at least the former side is still in development on the formal level. And to the extent that it has been developed, it ought to be uninteresting to see if, for example, Prawitz' account of logical consequence – that serves as a central impetus for proof-theoretic semantics (cf. ibid., introduction and sect. 1) – accords well with the work that is founded on it. Thus, it seems to me that a proper assessment of the LNLT is, for the time being, to proceed on a more conceptual level, as opposed to dealing with specific accounts. However, we will see in the following sections that we can *still* excavate a list of detailed criteria from the general ideas behind proof-theoretic accounts of logical consequence and constants. As such, our mission has by no means been rendered trivial, even if we sidestep technicalities.

For these reasons, I shall allow myself to be brief with the formal details, mentioning them only when necessary. Otherwise, I shall be focussed on the general ideas and motivations behind the proof-theoretic conceptions, and extract more conceptually-oriented realisability conditions as the end of this chapter.

4.2 The Proof-Theoretic Conception of Logical Constants

Coming directly from section 3.4, we have already met the basic idea of characterizing logical constants by drawing upon largely proof-theoretic means. The idea, given the background inferentialist framework, is to characterise logical constants in terms of rules of inference, which are thought to be their meaning (cf. also MacFarlane 2017, sect. 6). This, of course, does by itself not distinguish logical from non-logical vocabulary. For example, the introduction rule of a predicate like that of being a bachelor might be characterisable in inferential terms as follows:

$$\frac{x \text{ is a human male} \qquad x \text{ is of marriageable age} \qquad x \text{ is not married}}{x \text{ is a bachelor}}$$

This seems to be entirely in the same spirit as the (purported) rules for "and" as seen before, corresponding to the rules of conjunction. Thus, more needs to be said to demarcate logical vocabulary from non-logical one.

The basic idea might be to characterise logical constants specifically by having 'purely-inferential' rules, i.e. having rules that do not depend or turn on some specific substantive content. A clear example is conjunction with its three associated inference rules:

$$\frac{A \qquad B}{A \wedge B} \qquad\qquad \frac{A \wedge B}{A} \qquad\qquad \frac{A \wedge B}{B}$$

Of course, the challenge here is to spell out in precise terms what "purely inferential" is supposed to mean. As MacFarlane points out, we also have to make further decisions. First, we would have to decide whether we use natural deduction or sequent calculus rules. Second, we need to decide whether the introduction or the elimination rule(s) (or both) are to be given primacy concerning meaning-constitution (cf. ibid.).[40]

Concerning meaning-constitution via introduction or elimination rules, it should be clear coming from section 3.4 that we are committed to treating *both* as constitutive of meaning. Since both the grounds of assertion and its appropriate consequences count towards understanding an expression, we shall straightforwardly apply this idea to the realm of (formal) logic. Somebody who can only recognise the introduction rule(s) of a connective as valid, while failing to appreciate the appropriateness of its

[40]He lists three more points (cf. ibid.), but these will either be ruled out in what follows or have been ruled out before, such as in relation to quantifiers in chapter 1.

elimination rule(s) cannot be credited with having understood the relevant connective. There is no reason to suppose that the notion of linguistic understanding and its criteria lose their significance or relevance when switching from natural languages to more formal ones. Logicians, as users of a specialist language, fall under the providence of the concept of linguistic understanding just as much as ordinary, i.e. non-specialist, speakers of natural languages. Thus, we shall treat both introduction and elimination rules as constitutive of meaning.

As I have indicated in section 3.4 as well, the idea of assertion-grounds and appropriate consequences finds it most natural expression in natural deduction systems with the relevant introduction and elimination rules. In contrast to this explicit distinction between introduction and elimination rule, what we find in sequent calculi is a distinction between left- and right-handed introduction rules. While a sequent calculus usually has the added benefit of making structural assumptions about the proof system explicit in its structural rules, an inferentialist cannot opt for such a representation. The main reason is this. In order to square sequent calculi with ordinary inferential practice, we have to read the multiple-conclusion formalism of such a proof system disjunctively. For example, in order to make sense of the sequent $A, B, C \vdash A \wedge B, C$, we have to read the conclusion as $(A \wedge B) \vee C$. However, if the meaning of disjunction is supposed to be fully captured by its rules of inference, it seems that we implicitly rely on this understanding when spelling out the usual right-hand introduction rule for disjunction (cf. Troelstra & Schwichtenberg 2000, 61):

$$\frac{\Gamma \vdash \Delta, A_i}{\Gamma \vdash \Delta, A_1 \vee A_2}\,\mathrm{R}\vee_{1,2}$$

where the conclusion would have to be read as a sequent from Γ to the formula $(\bigvee \Delta) \vee (A_1 \vee A_2)$, presupposing an understanding of "$\vee$" (cf. Steinberger 2009, sect. 12.3 and 12.4, especially 202f.). As such, we can forgo sequent calculi in favour of seeing logical constants as having their meaning constituted by introduction and elimination rules as they appear in natural deduction systems.

Before returning to the issue of spelling out the notion of being 'purely inferential', there is, following MacFarlane (2017), another decision to be made. As he puts it, we can disambiguate "characterise" as it appears in "inference rules characterise the meaning of logical constants" in two ways, following a roughly Fregean terminology. On the one hand, characterisation can be read as the rules in question fixing the reference of the logical constants. For example, as Rumfitt (2000, 806f.) argues, his bilateral[41] proof system for classical logic allows one to recover the classical truth-tables for these connectives under the assumption of bivalence. On the other hand, we can understand these rules as fixing the sense of a connective, i.e. one can grasp the sense of connective once one has grasped its introduction and elimination rules (cf. MacFarlane 2017, sect. 6). Coming from section 3.4, it should be clear that by identifying the meaning of a connective with its rules of inference, specifically its introduction and elimination rules, we are doing both simultaneously. Rumfitt's argument compelled us to treat the sense as given by the relevant rules of inference, while our further identification of the meaning with rules of inference in general make the former task trivial.[42]

The idea of 'sense determination' naturally forces one to rule out examples such as "tonk", which shows that not *all* pairs of introduction and elimination rules generate a coherent meaning. As we have seen in our discussion of logical inferentialism, the natural way to counter this is to adopt harmony as a necessary condition for logicality. We have, however, following Rumfitt (2017), rejected harmony as a necessary criterion for logicality. As he points out, all that we really need is the notion of global conservativeness, if we wish to adhere to the principle of innocence. Otherwise, the following requirement for individual logical constants suffices to rule out examples like "tonk" either way. The so-called requirement of conservativeness is that a newly introduced connective c with its introduction and elimination rules should not allow us to derive c-free formulas that were previously underivable in the base system to which c was added (cf. Steinberger & Murzi 2017, 208). It should be clear that this kind of conservativeness is a relative property that depends on the previously given base system. As such, a logical constant might be seen as faulty given one base system, but not another. As Steinberger puts it:

> There would be nothing objectionable about *tonk* so long as we are careful to introduce it only into contexts where it does not do any damage. (Steinberger 2011, 625)

While this is certainly true, it seems hard to find such a context. In fact, the only

[41] A proof system is bilateral if it distinguishes the rejection of a formula from asserting its negation, and unilateral if it identifies the two, cf. Rumfitt (2000, 797).

[42] All of this, of course, to the degree that we even wish to be put into the dichotomy of reference and sense. Given our commitments in chapter 3, I suppose we should not, and even so, it is clear that under our commitments the two dimensions of meaning collapse into each other, at least for logical vocabulary.

context imaginable is a logic which has either no theorems or contains every single atomic proposition to begin with. For as long as we can derive so much as one theorem in our system, we can derive every well-formed formula. Hence, "tonk" is only innocent where we have nothing interesting to prove in the first place. Insofar as our ultimate goal in relation to the LNLT is concerned, then, this point is hollow anyway. If we were to add "tonk" into natural language as logical constant, triviality *would* ensue. Thus, with regards to *ordinary* inferential practice, "tonk" *is* vicious in virtue of trivialising an important aspect of human intellectual activity – deductive reasoning.

There is one last addition to be made before moving on to characterising purely inferential rules. We have said nothing so far about which I- and E-rules are to be seen as meaning-constitutive, as there is arguably some variety to choose from in virtually all cases. Luckily, the notion of primitive obviousness alleviates this issue, thus we shall further require that the I-and E-rules are each primitively obvious as specified in section 3.4.

Let us summarise the findings so far. We have committed ourselves to seeing the meaning-constituting rules of connectives to be given in the general style of natural deduction rules, where we decided to treat both I- and E-rules as constitutive of meaning. Furthermore, having rejected harmony as a criterion for logicality before, we shall at best require conservativeness for newly added connectives to any base system.[43] What we are left with now is to spell out the sense in which the rules characterising a logical constant are said to be 'purely inferential'.

Given our introductory examples above, it seems that what distinguishes the purely inferential rules such as those governing conjunction from those of the concept 'bachelor' are the content of the premisses and the conclusions. In a sense, in the case of 'bachelor', the specific content of the premisses mattered in order to establish the conclusion. We could not have established that x is a bachelor without the information that x is a human male. Moreover, the consequences we can draw from an assertion such as "x is a bachelor." are numerous, with arguably a large portion of them relying once more specifically on the concept itself. In contrast, with respect

[43]Since our goal will ultimately be to find logical constants in natural language, i.e. *pre-existing* ones, we will not be occupied with this property any longer. The quick discussion of conservativeness simply served the purpose of motivating the proof-theoretic conception in virtue of showing that there *are* remedies against 'tonkitis', and that as such, the very idea of a proof-theoretic conception of logical constants is not philosophically bankrupt to begin with.

to the rules of conjunction, no specific concepts seem to matter. Of course, this is nothing else but an expression of the idea that logical constants – by themselves or concerning the inferences they alone license – are 'topic-neutral'. However, for the inferentialist, this is just the idea that they are governed by purely inferential rules. Thus, we have to make a motivated decision about what we ought to count as purely inferential and what not.

We shall adopt the following straightforward characterisation. We require the relevant rules to only consist of two things: (i) *one* occurrence of the logical constant, either in the premisses or the conclusion,[44] and (ii) only schematic symbols otherwise, namely structural symbols, such as commas, or schematic letters as stand-ins for propositions (like the letter "A" as we have been employing it so far). If a purported logical constant's introduction and elimination rules satisfy these requirements, we can conclude that the logical constant is characterised by purely inferential rules. As such, the usual rules for conjunction, disjunction and implication are all verified as purely inferential immediately.

There is a slight issue with this proposal, however. As MacFarlane points out, these criteria are not satisfied by negation introduction, since it involves a second logical constant, namely "$\bot$":

$$\dfrac{\begin{array}{c}[A^i]\\ \mathcal{D}\\ \bot\end{array}}{\neg A}\,\neg\mathrm{I}_i$$

Surely, though, if anything is count as a logical constant, it ought to be negation.[45] As such, MacFarlane concludes, we need to either relax the condition for being purely inferential, or "[...] add more structure [.]" (MacFarlane 2017, sect. 6). This argument, however, solely relies on seeing "$\bot$" as a logical constant in its own right, which is something that has been contested (Rumfitt 2000, sect. IV, and Steinberger 2009, sect. 2.7). It seems, rather, that there are good reasons to see "$\bot$" as a 'deductive punctuation mark'.[46] One is typically told that "$\bot$" stands for

[44] This ought to account for the intuition that, for example, basic competence of conjunction is arguably independent of basic competence of disjunction, and vice versa. This obviously does not rule out the idea that *mastery* of the connectives *does* involve competence of their various interplays.

[45] Indeed, logicians' occupation with the correct principles involving this connective would be unintelligible otherwise.

[46] As such, "$\bot$"'s purported elimination rule – *ex falso quodlibet* – becomes a *structural* rule (Steinberger 2009, sect. 2.7). This is due to the fact that it (thus) only contains either schematic letters ("A") or structural symbols ("$\bot$").

"0 = 1", or some other necessary falsehood. As Rumfitt points out, however, with regards to ordinary inferential practice, this is strange at best. Consider the following mathematical argument:

> "Suppose there is a greatest natural number, let us call it N. However, by the Peano axioms, if N is a natural number, then so is its successor $S(N)$. Moreover, we have that $N < S(N)$ by way we defined "<". Contradiction. So, there cannot be a greatest natural number."

In its standard formalisation, this argument would finish with an instance of $\neg\mathrm{I}_i$, so we need to understand "$\perp$" in relation to the verbalised piece of argumentation above. Clearly, "$\perp$" ought to stand in for "Contradiction.". Yet is "Contradiction." making an assertion? And if so, with what content? As Rumfitt says: "[...] it would be perverse to try to assign a propositional content to the expression 'contradiction'." (Rumfitt 2000, 793f.). Thus, "$\perp$" indicates that we have reached a deductive dead-end, and nothing more.[47]

This way of spelling out the notion of being purely inferential is, of course, rather restrictive, hence we might rule out a lot of expressions as logical that would otherwise be taken be such. This should not worry us, though, as we intentionally set out to ignore anything that goes beyond the sentential connectives, and as we have argued above, this criterion *does* give us those. Moreover, any criterion that is strict would only serve to make a positive assessment of the LNLT harder, so this cannot be held against us in this regard either. Thus, we will stick to these criteria in what follows.

4.3 The Proof-Theoretic Conception of Logical Consequence

The general idea behind the proof-theoretic conception, as was to be expected, takes a different starting point than its model-theoretic cousin. Instead of the preservation

[47]One might lament the given formulation. Could we not have said "This is a contradiction." instead of "Contradiction.", the latter being a shorthand for the former? And if so, would this not imbue "$\perp$" with propositional content? I think not, because it is easy to construe the extended version as simply pointing to the fact that we have reached a deductive impasse. Taken literally, it seems to suggest that "contradiction" is a property, rather than a proposition. As such, we do not *need* to understand that sentence as imbuing "$\perp$" with propositional content. Moreover, equating "$\perp$" with "This is a contradiction." cannot work either way, for the latter is itself *not* a necessary falsehood.

of designated values, the notion of 'proof' takes center stage. The general idea can be stated as follows:

> An argument from a finite number of premisses A_1, ..., A_n to the conclusion A is valid if and only if there is a *proof* of A via the premisses A_1, ..., A_n.
> (cf. Beall, Restall & Sagi 2019, sect. 3.2)

This, of course, means that we need an account of what it means to be a proof of something, and on pain of circularity, this cannot involve the notion of logical consequence itself. As such, the mission is to explicate the notion of proof in a satisfactory way. We will see how this works further below.

First, note that one of the key aspects of this conception is its emphasis on the epistemic aspects of consequences (cf. ibid.), something which its model-theoretic cousin is not taken to account for. The usual argument runs as follows. Logical consequences can be used to expand our knowledge, for example in the sense that knowledge of the premisses and knowledge of the (logical) validity of the inference to be drawn justifies a claim to knowledge of the conclusion, in virtue of the validity of the inference guaranteeing the truth of the conclusion. However, according to the model-theoretic conception of logical consequence, establishing validity of an inference *already* presupposes knowledge that for the given premisses and the given conclusion, it is not the case that the premisses are true and the conclusion false. In other words, knowledge of validity already 'includes' the knowledge about the truth-value of the conclusion, and hence this conception of logical consequence is 'epistemically inert' (cf. Etchemendy 2008, 267, and Prawitz 2005, 675f.).[48][49]

In contrast, the proof-theoretic conception is concerned with the inference steps that are taken in order to establish the conclusion. To put it differently:

[48]Etchemendy raises further issues for the account apart from its general conceptual inadequacy, namely for its extensional adequacy as well. Cf. Etchemendy (2008, 271-282) for the details. Prawitz himself further notes that the purported modal ingredient in the model-theoretic explication might, after all, be hollow. Cf. Prawitz (2005, 676).

[49]As Luca Incurvati told me in private correspondence, some authors take issue with the basic assumption (or intuition) that logic is indeed supposed to be epistemically 'active' as opposed to inert in this sense. However, we already saw an example above coming from Rumfitt (2017), where an application of disjunctive syllogism yielded precisely such epistemic gain, hence I think the assumption is in good standing. Moreover, it seems strange to think that proving things, as a kind of doing, does not stand in the service of epistemic goals. Yet clearly this would square badly with logical consequence being epistemically inert.

> A proof does not merely attest to the validity of the argument, it provides the steps by which we can establish this validity. And so, if reasoner has grounds for the premises of an argument, and they infer the conclusion via a series of of applications of valid inference rules, they thereby obtain grounds for the conclusion.
> (Beall, Restall & Sagi 2019, sect. 3.2)

As such, the validity of the rules of inference has to be established by means other than truth-value preservation, if it is to serve our epistemic goals.

Arguably the most detailed account of a proof-theoretic conception of logical consequence can be found in the writings of Dag Prawitz (1985, 2005 & 2006). As I have noted in section 4.1, we have little impetus to delve into the technical details of his account, which is why I wish to focus on some of its conceptual cornerstones instead.

The basic idea of Prawitz' account is that certain rules of inference of certain expressions form the meaning of the latter, which in turn forms a 'rock-bottom' from which any further justification of inference steps or arguments proceeds. Prawitz observes that under certain circumstances, a sceptical questioning of an assertion or inference can be rejected on the grounds that the assertion or inference obtains in virtue of the meaning of the expressions involved (Prawitz 2005, 681f.). He considers the example of "$3 + 1 = 4$". If someone were to question this statement, it seems a natural answer is to point to the fact that the number 4 just is the successor of 3, as a matter of fact of what "4" means. Setting aside mathematics, he considers the following example in relation to inferring:

> Similarly, what can we answer someone who questions the drawing of the conclusion $A \to B$, given a proof of B from A, except that this is how $A \to B$ is used, it is part of what $A \to B$ means?
> (ibid., 682).

This is *the* fundamental idea behind proof-theoretic conceptions of logical consequence: Certain inferences are 'simply' valid, on the very basis of how the terms involved are used. Furthermore, all other inferences shall be justified with reference to those rules.

On the basis of this idea, Prawitz then builds up an account of logical consequence in terms of valid arguments. By an argument, we can think of something along the lines of proof trees as they appear in natural deduction systems (cf. Prawitz 2006, 511ff.). In general, arguments are valid provided they end in an application of one of our basic, meaning-constitutive rules of inference and their immediate subarguments are valid (cf. ibid., 512). Such arguments are said to be in canonical form. In case an argument is not already given in this way, we shall require specific procedures that transform such an argument into one that is canonical form (cf. ibid., 513f.). In the context of natural deduction systems, those procedures take the familiar form of normalisation algorithms (ibid., 514). On the basis of this, an argument is called "logically valid" just in case it is valid irrespective of the atomic propositions we started out with. In other words, an argument is logically valid just in case its validity does not hinge on both the content and the placement of atomic propositions within the argument (ibid., 515f.). This blueprint of how validity is to be understood is then also applied to individual inference rules. A rule of inference is (logically) valid if and only if there is a procedure that transforms any argument that ends with an application of said rule into one that is recognisably valid, provided its immediate subarguments are (irrespective of the involved atomic propositions) (ibid., 516).

As a matter of fact, we have already seen this idea in section 3.4. Recall the peak-levelling procedure for *modus ponens* and *conditional proof.* Since Prawitz takes the canonical argument for a conditional to be given via *conditional proof* alone (cf. ibid., 510), he provides the following justification for *modus ponens.* He invites us to assume we are given an argument of the following form:

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ A \rightarrow B\end{array}}{B}\,\text{MP}$$

with $\mathcal{D}_i$ being the immediate subarguments. Assuming those to be valid means that $\mathcal{D}_2$ can be brought into canonical form:[50]

$$\dfrac{\begin{array}{c}[A]^1\\ \mathcal{D}\\ B\end{array}}{A \rightarrow B}\,\text{CP}_1$$

[50]This stems from the inductive flavour of Prawitz' account of validity. By assuming these subarguments to be valid, the notion of transformation into canonical form applies to them per induction hypothesis (cf. ibid., 515).

Now, using the same trick as we did in section 3.4, we can conjoin these arguments to give us a valid argument to the same conclusion, but one which is recognisably valid, since we assumed $\mathcal{D}_2$ to be valid, hence its subargument – $\mathcal{D}$ – has to be as well:

$$\begin{array}{c} \mathcal{D}_1 \\ A \\ \mathcal{D} \\ B \end{array}$$

This, then, concludes the justification of *modus ponens* for Prawitz, since we can bring *any* argument ending in an instance of it into a form that is valid, provided its subarguments are (ibid., 516). In other words, *modus ponens* preserves validity.

However, we have shown even more. With this reasoning, we have furthermore established that *modus ponens* is logically valid. For observe that neither did the exhibited 'justification procedure' depend on the specific content of A, B or any atomic propositions, nor anything else in the above reasoning. As such, this reasoning applies universally, and we have established logical validity. This point is crucial, since without us having the means to establish logical validity in this generality, it seems doubtful that the proof-theoretic conception of logical consequence is not also rendered epistemically inert in its own way. If we had to check every set of atomic propositions individually to establish logical validity, it seems that knowledge of logical validity would be out of reach. As we have seen, however, the way to establish such validity is on the basis of reasoning with arbitrary arguments, which specifically *relied* on the fact that *modus ponens* is given in *purely inferential* terms already. In this way, we can account for the formality of logical consequence, i.e. – following Prawitz – the idea that validity is retained no matter any uniform substitution of atomic propositions (cf. Prawitz 2005, 672f.).[51] As such, the proof-theoretic conceptions of logical consequence relies on a prior demarcation of logical constants (cf. also Beall, Restall & Sagi 2019, sect. 3.2).

To round off our investigation into proof-theoretic conceptions of logical consequence, let us return to its epistemic-modal ingredient. The idea is that the meaning-constitutive rules of our connectives are valid in virtue of their very meaning. Nevertheless, traditionally conceived, such rules ought to preserve truth, and the question remains on how the proof-theoretic conception can account for this. We have encountered truth earlier in our discussion of Brandom, where he claimed

[51] We will revisit this point when discussing ordinary inferential practice in section 5.2.2.

that truth may serve as a tool for characterising good inferences. However, as far as ordinary assertion practice is concerned, we argued – with Brandom – that truth-conditions are superfluous, at least as an essential ingredient. Still, recall that our office-example did demonstrate that for some assertions, the grounds on which it would be appropriate to infer it, and hence what a defence of it would amount to, include verification conditions. Nonetheless, we shall construe validity in general as ultimately being tied to a preservation of *assertibility*. This would allow truth to still play a role to the extent that assertion conditions are tied to means of verification.[52] In those cases, the idea that valid arguments preserve truth is that in virtue of preserving such verification conditions, if we had conclusive evidence for the premisses, we are thereby in possession of conclusive evidence for the conclusion (Prawitz 2005, 681ff.). In general, though, what really counts is that assertibility or commitment is preserved, not truth (cf. also Beall, Restall & Sagi 2019, sect. 3.2).

How is assertibility preserved, exactly? Under the Brandomian picture, an inference, if appropriate given the linguistic community's standards, is such that if I am committed to the premisses, I am, given the appropriateness, also committed to its conclusion. But why? Recall that an assertion involves commitment, which means that an interlocutor must be prepared to defend their claim. This is done by providing further assertions, from which the original claim can be appropriately inferred. It is the premisses of such inferences that can be counted as the conditions under which an assertion of the conclusion is warranted, hence an appropriate inference links such conditions with a warrant for its conclusion, in virtue of being an appropriate inference. It is in this way, that rules of inference can preserve assertibility.

With these details in hand, we can now also appreciate the way the proof-theoretic conception accounts for the necessity involved in logical consequence. The usual proof-theoretic gloss speaks of a 'necessity of thought' with which the conclusion follows from the premisses. This necessity obtains, because the conclusion follows via primitively obvious rules of inference that constitute the very meaning of the relevant connectives. Hence, in virtue of participating in the pertinent linguistic practice, we are committed or compelled to accept the conclusion arrived at by such means, provided we are already committed to the premisses (cf. ibid.). In this

[52]This, I hope, accounts for the standard examples used to demonstrate the necessary truth-preserving character of logically valid inferences, since those typically involve empirically verifiable sentences.

way, the proof-theoretic conception accounts for both the formality and the necessity constraint on logical consequence.

4.4 Summary – The Realisability Conditions

It is time to list our realisability conditions. From what we have gathered so far, we can see that in order for there to be logical consequence at all, we must be able to first distinguish, within the language, between logical and non-logical vocabulary.[53] In order to count as logical, an expression must satisfy the following criteria:

(LC1) Its meaning can be given by introduction and elimination rules in the style of a natural deduction system.[54]

(LC2) These rules in question are primitively obvious, as characterised in section 3.4.

(LC3) The rules are purely inferential, in the sense that they contain only one occurrence of the (purported) logical constant in either the premisses or the conclusion, and only structural or schematic signs otherwise.

On the basis of such vocabulary, we can now specify what it means for a (finite) set of premisses Γ to logically entail a conclusion A:

(LCons) A is a logical consequence of Γ iff there is a logically valid argument from premisses in Γ to the conclusion A.

This characterisation is not yet complete, as we must further specify what a logically valid argument amounts to.

Coming from our discussion of Prawitzian ideas and our account of logical constants, the idea is that certain rules of inference can only be defended by reference to them being constitutive of the very meaning they give to the connective in question. As such, what we require is that in order for an argument to be logically valid, it must be in the form of a series of correct applications[55] of *only* such rules. In a more compact formulation, we get the following characterisation:

[53]Very much in the spirit of the argument from logical constants in chapter 2.

[54]The formulation with "can" serves the function of allowing the translation of an expression of an informal language into a more formal key, in order to avoid the trivial mismatches mentioned in chapter 1.

[55]Correct as laid out in the rules themselves, for example concerning order of premisses. Since we rely on natural-deduction-like rules, we can see correctness of application to be determined in like-wise fashion. This also ought to tie up the observations of Shoesmith and Smiley from section 3.4.

(Val) An argument from premisses Γ to the conclusion A is logically valid just in case it is a series of applications of either introduction or elimination rules of logical constants.

Thus, we have a logically valid argument just in case each inference step is gap-free and belongs to the meaning of a logical constant. This characterisation is not circular, despite the occurrence of "logical(ly)" on both sides in (LCons), as the logicality of the validity arises due to the basic rules of inference being tied to logical constants, themselves specified further without reference to anything itself logical.

In sum, this formulation pays heed to the epistemic element of logical consequence, given that a logically valid argument will be such that each inference step is analytic, in the sense that any doubts about its propriety can only be rejected with reference to its primitively obvious nature, i.e. that the rule is part and parcel of the meaning of the involved connective. Moreover, as we have argued above, it preserves assertoric commitment in virtue of its analyticity. Lastly, the formality constraint in logical consequence is obtained in virtue of those inference rules being purely schematic.

With regards to providing a coherent package in the same vein as the one Glanzberg chose as his starting point – model-theoretic conceptions and referentialist truth-conditional semantics – we can now see that we are in a similar position with regards to inferentialism and proof-theoretic conceptions. Our global inferentialist commitments lead us to logical inferentialism, which in turn informed and shaped our conception of logical constants. These, on the other hand, ended up as the central ingredients in the conception of logical consequence, hence providing a link between global inferentialism and logical consequence. With this alternative package in hand, we shall now turn to assessing the question of whether natural language contains logical consequence relations or not.

5 The LNLT Revisited: An Inferentialist Take

We are finally in a position to evaluate the logic in natural language thesis (LNLT) with respect to inferentialism and proof-theoretic conceptions of logical consequence. As we have seen, the primary ingredient in the proof-theoretic conception was the initial demarcation of logical and non-logical vocabulary, and it is with this distinction that we will start. This will also constitute the inferentialist response to Glanzberg, taking the argument from logical constants (ALC) as its starting point. Having dealt with the issue of logical constants in natural language, we will then move on to study the possibility of reconstructing logical consequence relations in natural language, on the basis of excavated logical constants. All this will, at least *in principle*, secure logical consequence relations in natural language. I will further argue that there are good reasons to be optimistic about the actual existence of logical consequence relations in natural language, on the basis of considering the pertinent example of "and". These considerations will provide us with the inferentialist's answer to both Glanzberg and our initial 'big-picture' puzzle – the mismatch between the methodology as well as epistemology of logic, and the purported non-existence of the relevant entities in natural language.

5.1 Logical Constants in Natural Language

Let us first recall the ALC. Glanzberg argued for two things: First, natural language does not come pre-equipped with a demarcation between logical and non-logical vocabulary, and second, the purported logical constants could only ever give rise to lexical entailments, given the absoluteness of natural language semantics. The first point has been explored under the idea that perhaps grammatical demarcations allow for an isolation of logical vocabulary. However, as Glanzberg argued, such criteria would arguably over-generate, trivialising the formality constraint on logical consequence. The second point is once more reliant on the absoluteness of natural language semantics. Given its adoption due to the requirement of accounting for linguistic understanding, the entailments that purported logical constants give rise to could only ever be *lexical*, i.e. on par with the entailments as provided by other expressions of natural language. As a last note, it should be clear that we cannot individuate logical constants on semantic grounds either, since that, according to the model-theoretic picture, would require permutation invariance – a criterion once more blocked by the absoluteness of natural language semantics.

Since the inferentialist can block the argument involving absolute semantics, we can safely ignore the issue concerning a grammatical distinction between logical and non-logical vocabulary. As far as the ALC is concerned, there seems to be a general mismatch involved, specifically between the requirements for logical constants as well as consequence relations, and the requirements for semantics in order to model linguistic understanding.[56] I now wish to comment on this mismatch a bit further, hence we shall take a step back and consider the ALC in a more abstract light.

The principal problem seems to be what I would like to call a 'conceptual mismatch'. On the one hand, the model-theoretic conceptions require certain tools to be brought to table in order to get working, while on the other hand, tying linguistic understanding into semantic theorising deprives us of precisely those tools. As such, it seems there is a certain mismatch on the *conceptual* level: The necessary conditions for logicality – be they for constants or consequence relations – according to model theory are simply absent in the conception of linguistic understanding and meaning that is at the core of the referentialist truth-conditional picture, as advocated by Glanzberg. Thus, it is, *in principle*, not possible to have logical constants or logical consequence relations in natural language. We have reinforced this last point further in chapter 2 when discussing Sagi, finding her approach to rescuing the LNLT wanting, again on grounds related to the possibility of accounting for linguistic understanding.

We are now in a position to see a distinction between two kinds of ways in which the LNLT could fail: on the conceptual or the extensional level. The former relates to a mismatch between different conceptions or concepts, while the latter, assuming those conceptions or concepts to be satisfiable in principle, considers whether we can actually find anything in natural language that satisfies them. As we have seen, the model-theoretic conceptions suffer a conceptual mismatch, rendering any extensional search vacuous to begin with. In what follows, I will scrutinise proof-theoretic conceptions along both lines as well, and argue that there is no conceptual mismatch, and reasons to be optimistic concerning the question of extensional 'satisfiability'. Let us first turn to the conceptual level.

[56] Recall its formalised version from section 2.1.4.

5.1.1 A Conceptual Mismatch?

It should be clear, coming from chapter 3, that the inferentialist takes the relationship between meaning and understanding serious, and any ascription condition for logical notions that would take us beyond the realm of what is needed to account for linguistic understanding would clearly be as catastrophic for the inferentialist and their take on the LNLT as it was in Glanzberg's case. Since it was the quest for an account of linguistic understanding that introduced absoluteness and generated the conceptual mismatch, I shall first discuss each ascription condition for logical constants under the proof-theoretic prism with regards to the issue of going beyond linguistic understanding, before widening the scope and considering inferentialist (meta)semantics in general.

Let us start with primitive obviousness (LC2). Recall that a rule of inference is primitively obvious just in case it is (i) obvious, and said obviousness is primitive in a specific sense (ii). Given the way we analysed (i), it is only concerned with actual speakers' linguistic behaviour. For a rule of inference counts as obvious just in case a speaker is disposed to assent, affirm or simply not object to the rule whenever it is instantiated, without any need for further elaboration or justification. Under this understanding of "obvious", could it be that in order to satisfy this condition, a candidate logical constant could take us beyond the realm of linguistic understanding? The answer is a clear "no". For the notion is grounded in the linguistic behaviour of speakers themselves. In order to satisfy this condition, we require a certain attitude towards the rule – there is nothing in principle about our meaning-theoretic assumptions that would rule out this possibility *a priori*. We simply require of the rule that it has a certain epistemic profile, but this clearly does not take us beyond the *rules themselves*, which we decided to treat as the *explanans* of linguistic understanding, coming from the constitutive argument.

There is, moreover, no doubt about there being such obvious rules in natural language. Consider the classic example of inferring that someone is unmarried given that they are a bachelor, or the inference that if x is left of y, then y is right of x. These are classic examples of rules of inference that can be found in natural language, which are also primitively obvious. Inferring from "x is left of y" that y is right of x is part and parcel of understanding the two relevant expressions, and no one who understands them would object to the inference under consideration. Moreover, the appreciation of the inference by a competent speaker is clearly not grounded in any prior grasp of other concepts, or by applications of further inference rules. The inference is as gap-free as they come. Thus, it seems we find primitively obvious rules in natural language anyway.

With this lesson in hand, we can immediately see that the primitiveness (ii) in question is equally in good standing. All of the criteria for primitiveness once more rely on speakers' linguistic behaviour, specifically the fact that their acceptance of the candidate rule is not grounded in the acceptance of other rules, iterated applications thereof, or prior grasp of beliefs independent of grasping the rule in question. Again, these criteria do not take us beyond the realm of linguistic understanding, i.e. the realm of rules of inference, since they simply impose further conditions on the rules themselves. In addition to this, such rules can already be found in natural language, as the two aforementioned examples concerning bachelors and left-right-directions demonstrated.

It seems, then, that as far as inferentialist metasemantics are concerned, there is no mismatch between the former and (LC2), for we are not forced to require anything besides certain rules of inference, albeit with a certain epistemic profile, itself cashed out in terms of the linguistic behaviour of speakers. Moreover, primitive obviousness is clearly not a feature of any 'lexical entries' we might have for our logical constants, but a means to determine which rules associated to a purported constant count as meaning-constitutive. As such, the criterion does not infringe on inferentialist semantics *per se*, and there is nothing about our metasemantics that prevents rules of inference with a certain epistemic profile to determine meaning, much less account for linguistic understanding. In fact, if the commitments would do so, the inferentialist would have a hard time accounting for any analytic inferences, as the previous examples involving bachelors and directions demonstrated. For those inferences, too, were primitively obvious, and, moreover, are usually taken to be meaning-constitutive for the expressions "bachelor" and "left" in inferentialist semantics, respectively.

In sum, (LC2) does not take us beyond the rules themselves, which we took to be the entities accounting for linguistic understanding, nor does it go beyond what the inferentialist is committed to in relation to their general metasemantics and semantics, either.

Consider (LC1) next. A logical constant's meaning-constitutive rules, among other things those that are primitively obvious, should have to be able to be given in the form of introduction and eliminations rules, akin to those of a natural deduction system. Under the assumption of inferentialism itself, especially the two-aspect model of meaning, *every* sentence has grounds for its warranted assertion, and appropriate consequences that can be drawn from it, which determine and constitute its meaning. As such, we *expect* there to be such grounds and consequences for any logically complex sentence, mediated by the (purported) logical constant occurring in it. In other words, the very idea of inferentialist (meta)semantics sets up such introduction and elimination rules in the first place. This leaves us only with the issue concerning presentation. Yet again, it seems that none of our (meta)semantical assumptions rule out that some rules can be given in the form of

$$\frac{P_1, \qquad P_2 \qquad \ldots \qquad P_n}{C}$$

and whether we can actually find such a presentation for our purported constants is a matter of extensional, not conceptual, satisfiability, which I shall postpone for later. In addition to this, we once more find 'companions in guilt' in natural language. Recall the rule that allows one to assert "x is a bachelor.":

$$\frac{x \text{ is a human male} \qquad x \text{ is of marriageable age} \qquad x \text{ is not married}}{x \text{ is a bachelor}}$$

Since there appears to be nothing above and beyond the given premisses to establish that someone is a bachelor, we have found yet another comparable example already present in natural language, in this case a rule of inference that allows for treatment in terms of an introduction rule as it would be encountered in natural deduction systems.

In the same sense that (LC2) did not really infringe on semantics, (LC1) does not really infringe on metasemantics. It is purely concerned with how certain meaning-constitutive rules are to be given as far as semantics are concerned. Those rules do, of course, also function as determiners of meaning on the metasemantical level, but nothing about that role is tied to a specific style of presentation. In any case, (LC1) is a requirement for semantic entries first and foremost, and is once more concerned only with rules of inference, hence does not take us beyond the inferentialist framework for natural language (meta)semantics. Thus, I submit, (LC1) faces no challenges on the conceptual level either.

Since the general nature of introduction and elimination rules are clearly not at odds with our inferentialist (meta)semantics, the question now shifts to the conceptual satisfiability of them being purely inferential (LC3). However, as we can see from our discussion of (LC1), (LC3) does not require us to go beyond inferentialist (meta)semantics for natural languages either, since (LC3) is again concerned only with a requirement *on rules*. We did not incur any commitment in chapter 3 to the extent that rules of inference could not be purely inferential, whether as determiners of meaning, as the *explanans* of linguistic understanding, or as the meaning of an expression itself. Whether we can actually find such rules in natural language is another question, namely one for extensional satisfiability.

Either way, what we can see here is a decisive difference between the model-theoretic and proof-theoretic conceptions. Whereas permutation invariance would have taken us beyond the referentialist framework accounting for meaning *and* linguistic understanding, the proof-theoretic conception faces no such issue. For the proof-theoretic way of conceptualising topic-neutrality is not permutation invariance, but the purely formal character of the relevant inference rules, which in turn does not take us beyond the realm of rules, but simply imposes further criteria on them.

Thus, I hope to have shown that the proof-theoretic conception of logical constants does not face the same kind of conceptual mismatch with an inferentialist take on meaning in the same vein as the model-theoretic variety mismatched with the referentialist's approach. None of the three criteria we took to be necessary for logical constanthood were seen to square badly with the inferentialist framework, but simply impose further requirements on the material that the inferentialist works with, namely certain epistemic and formal profiles for certain rules of inference. However, we have also seen that while none of the conditions individually caused a mismatch on the conceptual level, the question remains whether we can actually find any expression in natural language that satisfies all three and thus counts as a genuine logical constant.

5.1.2 The Example of "and"

As far as answering such sceptical doubts concerning the existence of expressions satisfying (LC1-3) are concerned, it will arguably be most instructive to consider a concrete example of a purported logical constant. This will also serve to demonstrate that even if there is no mismatch on the conceptual level, the inferentialist has their work cut out for them on the extensional level, which will be shown to be less than trivial. The increased expressive power of the inferentialist approach over the truth-functional one – recall Rumfitt's argument in section 3.4 – comes at the price of a need to account for the various inferences certain words license, specifically trying to account for them in terms of purely inferential, primitively obvious introduction and elimination rules.

The word we shall be considering is "and". The reason for this particular choice is two-fold. First, it is a natural candidate for logical constanthood, with rules of inference that allow for easier handling than others. Second, and in contrast to the first point, it will be shown that even in a case where the corresponding rules are generally taken for granted, the complexity of the ordinary use of "and" make this a more tantalising endeavour than what might have been anticipated. Still, I hope to show that even though accounting for the English "and" is complex, at the end of the day it is justified to treat "and" as a proper logical constant of natural language.

Recall our earlier rules for "and" as expounded in section 3.4, where we conjectured that those rules must be readily accepted in order for somebody to count as having understood "and":

$$\frac{\text{A} \qquad \text{B}}{\text{A and B}}\ \text{and-I} \qquad\qquad \frac{\text{A and B}}{\text{A}}\ \text{and-E} \qquad\qquad \frac{\text{A and B}}{\text{B}}\ \text{and-E}$$

Let us first look at whether these rules, and hence "and" by extension, satisfy (LC1-3). Clearly, both (LC1) and (LC3) are satisfied: The rules are given in the style of a natural deduction system, involve one introduction and two elimination rules, and further involve only one occurrence of "and" each, together with only schematic letters otherwise. Moreover, it seems feasible to see these rules as constitutive of the meaning of "and" under inferentialist lights, as not being able to appreciate their felicity arguably bars one from being accredited with understanding "and" to begin with. As Boghossian puts it:

> [I]t's hard to see what else could constitute meaning *conjunction* by 'and' except being prepared to use it according to some rules and not others (most plausibly, the standard introduction and elimination rules for 'and').
> (Boghossian 2011, 493)

In particular, then, somebody reasoning from "A" to "A and B", or from "A and B" to "It is not the case that A", *could* not be credited with having understood "and", and they would be corrected accordingly. Moreover, if what is understood when grasping those rules is the meaning of “and”, it seems we have a good case for taking the meaning of “and” to be constituted by and-I and and-E, as argued for in section 3.2.

What about (LC2)? It seems fair to say that these rules are primitively obvious, but let us tackle this in more detail. We remarked in section 3.4 that often, a rule’s being obvious coincides with an application of it going un-objected in ordinary inferential practice. Now, there is clearly nothing objectionable about me being told that Jane is waiting outside and Thomas having gone to the bathroom, and then answering with “Jane is waiting outside and Thomas went to the bathroom.” when asked where the two are, nor is there anything objectionable to the reverse order of inference, with me being told the whereabouts of both, and then relaying the information about either one of them to another interlocutor. The important part about obviousness, however, is not just the readiness of accepting the salient inferences, but for speakers to not require further justification or elaboration for them. Indeed, it seems impossible to justify or elaborate on this inference other than, at best, falling back on an awkward “Well, that’s just what “and” means?” or “Well, if you got A and B, you have A in particular, no?” when prompted to do so. Thus, it seems that those rules are obvious to competent speakers of English. In order to fully satisfy (LC2), however, such obviousness must be of the right kind. Yet as we have just seen, the reason why the call for elaboration or justification falls flat for such inferences is not because we can give a ready answer as to why the inference holds, but because there seems to be nothing about it that allows further justification or elaboration. By extension, thus, we cannot make use of other rules, repeated applications thereof, or need to use further beliefs to justify our drawing of the above inferences. From this, we can conclude that our initial rules satisfy (LC2) as well.

It would seem, then, that “and” is a *bonafide* logical constant, and this would be true, if those rules really were to exhaust the ‘core inferences’, i.e. the primitively obvious ones, associated with “and”, and if those rules were also licensed in exactly the generality that the rules seem to proclaim, namely in *all* generality.

It is not easy to come by any author these days that would have something negative to say about the comparison between “and” and “$\wedge$”, the latter whose rules we appropriated for “and” to begin with, given the widespread agreement that “and” and “$\wedge$” are close relatives, if not the same person.[57] P. F. Strawson, in his *Intro-*

[57] In this regard, recall the observations stemming from formal semantics as mentioned in chapter 1.

duction to Logical Theory (1952), however, offers reasons to be pessimistic about the idea that the rules given above match those of ordinary inferential practice related to "and". Despite the age of the work, it remains as one of the few in-depth systematic surveys of the relationship between natural language and formal logic, where the former is not already understood in a formal way that makes use of the resources of the latter. In other words, Strawson's work offers an unimpregnated view on the various uses of the word "and", free from assumptions stemming from work in formal semantics, or drawing upon a semantics-pragmatics divide.

This last point deserves emphasis. It has the added benefit of forestalling any objections in relation to the LNLT, in the sense that we cannot be charged with 'sneaking' the LNLT in through the back door by *imposing* our preferred rules onto natural language, then trying to gloss over any resulting mismatch with other explanations. In other words, it forces us to meet the inferential practice of natural language without any presumptions on the relation between the rules of formal logic and the behaviour of relevant words, such as "and". This way, we can be sure not to have incurred commitment to the LNLT without realising it, thus settling the worries raised in chapter 3 for good, as we are now taking a proper step back and forced to evaluate natural language without prior commitments.

What, then, are the reasons to suppose that our initial rules may not match the ordinary use of "and"? Strawson raises two issues. The first is that "and" allows for coordination effects, such as in "Tom and William arrived.", which is not something that "$\wedge$" can do in the usual formal syntax. Indeed, it does not seem simple to conceive of a formal notation to account for such coordination effects. Moreover, we cannot simply treat the aforementioned sentence as an abbreviation for "Tom has arrived and William has arrived.", since we would not say that people are speaking elliptically or in abbreviations when uttering such sentences (Strawson 1952, 79f.), to which I agree.

Despite my agreement with Strawson in this regard, there is still the fact that we *can*, validly, infer from such a coordinated sentence that Tom has arrived, or we *can*, equally appropriately, infer that William has arrived. In other words, it seems that and-E is preserved. But this is too quick. After all, in the case of coordinated sentences, the premise of and-E is not of the appropriate form. We would, thus, at least need to formulate further versions of and-E, and hence arguably also and-I, for coordinated sentences, though it might be a challenge to preserve (LC3) in this case. For example, one might be tempted to formulate and-E in the following way:

$$\frac{\text{and(A,B)}}{\text{A}} \qquad\qquad \frac{\text{and(A,B)}}{\text{B}}$$

One immediate issue with this proposal is that in order not to trivialise our own formality constraint (LC3), we cannot just introduce more formal notation. For, plausibly, there is a big range of expressions whose associated inference rules could be brought into a purely inferential key if only we were to do enough formalisation work beforehand. For example, if $\Box_x A$ is taken as a shorthand for "Agent x knows that A", then the following rule is purely inferential, even though it fails (LC3) without such prior formalisation:

$$\frac{\Box_x A}{A}$$

As such, we cannot just introduce more formal notation without trivialising (LC3). That said, there seems to be a big difference between allowing for more abstract notation in order to account for what are essentially only syntactic differences involving the *same* expression, and the formalisation of substantive expressions themselves.

In any case, we have multiple routes to choose here. We could provide a formal notation for coordinated sentences and treat all resulting rules as constitutive of the meaning of "and", given that they all satisfy (LC1-3), or we could try to give only one set of introduction and elimination rules, which are themselves stated generally enough to incorporate the coordinated sentences. As another alternative, we could treat "and" as different in meaning depending on whether it appears in a coordinated sentence or not.[58] In the latter case, we could maintain that this "and" *does* obey our initial rules.

This brings us to Strawson's second point. He observes that "Fx and Fy" does not always mean the same thing as "Fy and Fx". If I said that "They married and they had a child.", I would not be taken to say the same thing as when uttering "They had a child and they married.", because the "and" in this context does not merely 'glue' two propositions together, but also indicates a temporal order. This is not captured by the truth-table for "$\wedge$" – which is the account that Strawson is scrutinising – but neither by our initial rules (cf. ibid., 80).

This is, of course, no issue for and-E, which remains unchallenged. In turn, however, it seems that Strawson's example seems to suggest is that and-I cannot

[58]Cf. Heim & Kratzer (1998, 182f.) for the same observation in the context of orthodox semantics.

always work as we proposed, for the order of the sentences involved sometimes matters. Observe that with our initial rules, we are indeed beholden to treat “and” as commutative in this sense:

$$\dfrac{\dfrac{\text{A and B}}{\text{B}} \qquad \dfrac{\text{A and B}}{\text{A}}}{\text{B and A}}$$

and similarly for the other direction. However, in this case, the more expressive nature of the inferential account allows us to introduce more structure into our rules, ruling out the above ‘proof’ on the grounds that it does not respect the order of A and B in “A and B”. For example, we could require that the left to right direction of how the premisses are entered in the rule

$$\dfrac{\text{A} \qquad \text{B}}{\text{A and B}}$$

is to correspond to the contextually salient order between A and B. But we might not even have to go this far. It seems that the inferential profile of the two sentences above are such that we could equally well have said “They married, *then* they had a child.” and “They had a child, *then* they married.”. In each case, we could infer these new sentences from the old ones per assumption, given the specific context where “and” is used to imply a temporal order, while these statements clearly allow for inferences to the “and”-variants. In other words, it seems that the respective pairs of sentences are saying the same thing in such contexts. I would submit that this shows that the semantic work done by “and” in such cases is not *essential* to its meaning. Thus, from such inferentialist considerations *alone*, i.e. without reliance on any divide between semantics and pragmatics, we can see that such effects are not part and parcel of the meaning of “and”, and we can outsource the explanation as to why “and” implies some order in such examples outside of its ‘core meaning’, the latter which we can still capture using some variants of and-I as well as and-E.

Let us reconsider our situation. Strawson gave us some *prima facie* difficulties in simply assuming that the core rules for “and” are those of conjunction. As we could glean from the discussion, however, those examples are not a threat. Even if were were to fail to account for coordination effects without trivialising (LC3), we can restrict our attention to the set of sentences of the appropriate syntactic form, given that Strawson’s purported counter-examples for those instances were just that – purported.

Unfortunately, we are not out of the woods yet. As I claimed above, it is not easy to come by a critic of the assimilation of "and" with "$\wedge$" these days – but not impossible, either. There are more recent discussions concerning other linguistic phenomena that seem to impede on seeing the meaning-constitutive rules of "and" to be – more or less – those of conjunction. For a starter, anaphora seem to complicate the meaning of "and"-sentences. Consider the following one:

"The piano was brought downstairs, and it was loaded onto the truck."

with the "it" referring, anaphorically, to the piano in question.[59] Applying and-E would now result in the following conclusion:

($*$) "It was loaded onto the truck."

Now, one might object to this inference by pointing to the fact that the latter sentence by itself does not give the correct reading, since "it" by itself does not pertain to any specific individual, much less any pianos. Of course, what this objection ignores is the fact that within the context, "it" *is* specified. As such, if Clara tells Danielle that the piano was brought downstairs and it was loaded onto the truck, then it seems fine for Danielle to say "Okay, if *it* has been loaded onto the truck, I don't need to worry about *it* anymore!", and then to proceed not to worry about it anymore, since she inferred ($*$) to begin with. In other words, if the context supplies "it" with a specific reference, the inference seems licensed nevertheless.

There are more worries, however. While the inference to ($*$) may have been correct after all, there is still the issue that the following sentence is not giving the same reading as the original above:

"It was loaded onto the truck and the piano was brought downstairs."

As such, it seems that the order matters once more.[60] As I have remarked above, there might be ways to account for such effects by adding more structure to and-I. Even so, the question once more remains if "and" is the true culprit. For it seems clear that what causes the issues with and-I and and-E are not the rules themselves, but rather the occurrence of anaphoric phenomena. Perhaps we should understand

[59]Inspired by the examples in the literature on dynamic semantics, cf. van Eijck et al. (2016, sect. 2.1).

[60]Cf. also Dekker (2011, 930).

the premise in and-E as requiring that A and B are two sentences not linked by anaphora or other semantic devices, or perhaps my reasoning above is sound, and given a specific context, the inferences via and-E continue to be licensed. As far as and-I is concerned, we could equally restrict their application to semantically unlinked sentences A and B, hence guaranteeing universal applicability for those contexts, at least. A last piece of evidence pointing way from "and" as the culprit is that the central validities of dynamic predicate logic, so-called *Donkey-equivalences* (cf. Dekker 2011, 931), for example:

$$\exists x(\phi \wedge \psi) \equiv (\exists x\phi) \wedge \psi$$

crucially involve existential quantification, hence the dynamic effects seem to arise in conjunction[61] with update-effects, not "and" by itself. This is further acknowledged in dynamic predicate logic itself, where conjunction itself does not change the context, but remains sensitive to order of conjuncts (cf. ibid., 930).

As a last remark, we might have to disambiguate different uses of the word "and" either way. For example, in its use in the sentence below, it does not even seem to be employed as a means to glue different assertions together:

"A, B and C surrounded the fort."

For the inferences to "X surrounded the fort.", with X being one of the three, is only licensed if the conclusion is to be understood as "X [among others] surrounded the fort.", i.e. as X participating in a surrounding manoeuvre, since no one can surround a fort by themselves.[62] If one is happy to concede that this is, again given the context, an available reading, and-E might be upheld, after all. I would be unsure of this, however, especially given that the standard formalisation of such sentences is by use of so-called *plural quantification*[63] – a clear indicator we are not dealing with a 'distributive' use of "and", but with a 'collective' one (to steal some terminology from the literature on plural logic), i.e. the use of "and" is not to fuse different assertions together, but of another kind. If that is so, it stands to reason that we do not need to worry about such examples, for "and" is used in a different, albeit

[61] Pun not intended.

[62] Of course, we do speak of certain generals having surrounded certain forts, though arguably the reading in play is that the general *and their forces* surrounded the fort, or that the general's *commands led to* the surrounding of the fort.

[63] Cf. Linnebo (2004, 2017), for critical introductions.

similar if not analogous, capacity here – not to conjoin assertions, but *individuals*, as it were.

It would seem, this time with more confidence and reflection, that we have found at least one logical constant in natural language, namely the English "and".[64] With this, we have not only countered the ALC on the conceptual level, by showcasing how the proof-theoretic conception does not fundamentally mismatch in the way that its model-theoretic cousin does, but also on the extensional level, by giving reasons to believe that there is at least one actual logical constant in natural language. With this in hand, we can now move on towards logical consequence itself.

5.2 From Logical Constants to Logical Consequence

5.2.1 Ordinary Inferential Practice: What It Means to Follow From Premisses

Before we move on to check both the conceptual and extensional satisfiability of (LCons) with respect to natural language, it is necessary to get an overview of the ordinary inferential practice in question. As we shall see, there is at least one aspect to it that might be seen to square badly with (LCons). Nevertheless, I shall argue that the proof-theoretic accounts as unearthed in chapter 4 can deal with said aspect as well, which will in turn set the stage for a negative answer to the question regarding a potential conceptual mismatch.

I shall approach my characterisation of our ordinary inferential practice in the same spirit as Brandom's account of assertion, i.e. by considering it as a kind of public linguistic doing that ultimately answers to the community's conventions concerning appropriateness and warrant.

The first ingredient to our ordinary inferential practice is, of course, the drawing of inferences themselves. In general, an inference can be drawn either publicly or privately. In the former case, be it in the case of speech or text, the conclusion that was drawn is usually indicated by expressions such as "thus", "hence", "so", or "as such", appropriate intonations or the given sequence of utterances. In the private case, again given our assumptions about the relations between language and

[64]Plausibly, the same lessons can be applied to the Dutch "en", the French (and Latin) "et", the Spanish "y", the Italian "e(d)", the Scottish-Gaelic "agus", the German "und", the Chinese "和" etc. *ad nauseum.*

thought, it stands to reason that it functions similarly in the case of silent utterances made 'in one's head'.[65] Clearly, the temporal order with which utterances are made are not a fool-proof guide to what counts as conclusion and premise. If I say that the shovel should be in the shed or on the front porch, and you find it in the shed, it seems unobjectionable for you to reason as follows: "*Turns out* you were right, it was in the shed!". If we were to reconstruct the inferences made by, taking your indication that my being right is your conclusion at face value, a natural take is that you finding the shovel in the shed confirmed my statement concerning its location, which you (re)affirm with your saying of "You were right". As such, the premise is that the shovel was in the shed, and the conclusion is that it is in the shed or on the front porch, hence the conclusion does not always temporally follow the premisses. Moreover, this example showcases that not all premisses are always publicly uttered. As such, the drawing of inferences is a kind of linguistic doing that proceeds from premisses to some conclusion, which in general need neither be given in the order from premisses to conclusion, nor does either part need to be explicitly stated. This itself does not incur some strange scepticism about what counts as a conclusion and what as a premise, for there are ordinary methods of uncovering what an interlocutor intended as their conclusion and what as their premisses. Most straightforwardly, we do so by asking them, and the answer is given by their sincere avowals, usually by means of "if"-"then" constructions, which allow for easy individuation of premisses and conclusions.

The other two ingredients of ordinary inferential practice are now concerned with the case where things go sour. If the premisses and conclusions are clear enough for all participants and the inferences made are accepted by all of them, the drawing of the relevant inferences remains uneventful. In case this set-up is not established, two things come into play: the rejection of inferences and their defence.

As Dummett observes, a common way of subjecting inferences to criticism is to produce counter-examples. We guess what kind of inference pattern is at stake, and try to find a counter-example that is of the same form, but whose validity is clearly not granted, if not already taken to fail. This is usually done by using phrases such as "You might as well say [...]" or "That would be like saying [...]" (cf. Dummett 1991, 189). This also accords well with the use of the word "follow" in such contexts. If you object to my inference from every virtue having a corresponding vice to there

[65] For a nice take on this expression, cf. Ryle (1949, 35-40).

being a vice that is a counterpoint to all virtues, then you might say "This doesn't follow. Just because every natural number has a successor, doesn't mean there is a number that is the successor of all of them!". Coming from chapter 4, it might be thought that this aspect of ordinary inferential practice squares badly with (LCons) and (Val), given that validity cannot be established by considering individual counterexamples, on pain of rendering the proof-theoretic conception epistemically inert in its own way. I will return to this shortly. For the time being, we note that producing counter-examples is a central way of subjecting inferences to criticism.

Confronted with criticism, the original interlocutor can respond in various ways. They might withdraw their conclusion, on the grounds of accepting that their inference was invalid, they might add further premisses or even further inferences to strengthen the original conclusion, or they might object to the pattern as having been misidentified. In the latter case, they might respond with: "No, that's not what I had in mind, I meant to say [...]". Apart from the first case, all following moves in the game of inferring will lead to an increase of the number of inference steps performed, multiplying the number of conclusions and premisses accordingly. However, this process does not go on forever. Discarding pragmatic reasons, such agreeing to disagree or exhaustion, a 'cognitive' endpoint is reached when all premisses have been established, all conclusions identified, and all inference patterns have been determined. It is then that we have reached a kind of 'rock-bottom', in which all participants can clearly navigate the argument and reject or defend the contextually 'atomic' inferences in play. The ways in which this can proceed are once more numerous. It might be that one inference step enjoys insufficient support in the case of inductive reasoning, or somebody has put forth an inference as analytic given its constituent expressions, yet misunderstood them. Last, but not least, there are some inference steps which cannot be rejected whatsoever, because they hold 'by default', as it were.[66] Prime candidates would be proper analytic inferences – including the primitively obvious, purely inferential kind.

This last point ought to demonstrate that not only do logically valid inference rules tied to logical constants serve a useful expressive role in analysing arguments – as in the case of "if"-"then" constructions – but they provide a common founda-

[66]This picture is clearly inspired by dialectical approaches to logic. For entry points concerning the normativity of logic or the specific rules of inference such as *reductio*, cf. Dutilh Novaes (2015) and (2016). The inspiration comes naturally given the latter's attention to the actual doings of reasoners. That said, I do not share the genealogical underpinnings of her account, nor am I focussing on reasoning in the context of mathematics, specifically.

tion on which any reasoner can fall back on. As justifications have to come to an end somewhere – we would not call something a "justification" if it went on forever, discounting it as inherently incomplete – logical rules of inference provide just such a final destination for justifications in the game of inferring, while serving the additional function of expressive tools, thanks to their universal applicability.

Thus, overall, ordinary inferential practice consists in the drawing of inferences, which allows for complex interactions in terms of criticising and defending inference steps. In these cases, the process can proceed to a point where we reach 'rock-bottom', laying bare all relevant premisses, conclusions and inference rules. In order for this practice to be viable, however, there must be common grounds with respect to *some* inference rules, among which specifically logically valid ones are prime candidates. In the end, then, a conclusion follows from the given premisses if and only if the inferences establishing the former accord with the standards as set by the linguistic community,[67] something that can be revealed by playing the game of reasoning to the end, bottoming out all relevant inferences into their smallest intermediate steps. Those smallest steps will arguably be more often than not analytic inferences, including the logical kind – or so I tried to make plausible.

5.2.2 Conceptual and Extensional Satisfiability

Recall that according to our proof-theoretic conception of logical consequence (LCons), such a relation obtains between a finite set of premisses Γ and a conclusion A just in case we have a logically valid argument – or proof – from the premisses in Γ to A. Furthermore, an argument is logically valid if and only if it is an appropriate sequence of introduction or elimination rules constitutive of the meaning of some logical constants, where the appropriateness of a sequence is determined by the application conditions of the rules themselves. As I have noted earlier, and as we see here again, we cannot have a logical consequence relation without logical constants. However, once we do have such constants, certain inferences will automatically become logical consequences, namely those corresponding to single-step inferences of the relevant introduction or elimination rules. In other words, it seems that if we

[67] After all, the entire game as specified so far could only ever answer to the standards of said community, because under the Brandomian picture, those are the only ones needed to account for the appropriateness of inferences – be they tied to specific expressions *or* to reasoning itself, such as principles of inductive or deductive reasoning.

have at least one logical constant, we have a logical consequence relation, generated by the meaning-constitutive inference rules of the former.

As I have indicated with the last ingredient in characterising our ordinary inferential practice, logically valid rules ought to sit comfortably besides other analytic inferences. In other words, as far as the defensive aspect of our inferential practices are concerned, (LCons) not only finds a natural embedding, but the practice itself seems to rely on the existence of arguments *such as* those provided by (LCons). This does not absolve (LCons) – or by extension (Val) – yet, however, because the way we ordinarily criticise inference steps is by means of counter-examples. Since we cannot allow validity to be established on the basis of checking individual cases, it might be thought that our conception of logical consequence, given via (LCons) and (Val), does not allow for such counter-examples, and hence is conceptually inadequate with regards to our ordinary inferential practice.

This is not true, however. By the way that Prawitz has proven the logical validity of *modus ponens*, we saw that he relied on the fact that the rule itself was given in purely inferential terms. This negated the need to check for individual atomic bases or the specific placements of individual sentences in the argument. In our more general account, following Prawitz' lead, both introduction and elimination rules constituting the meaning of logical constants are valid by the very meaning they bestow upon the connective. In other words, they *cannot* have counter-examples, given both their analytic validity *and* their purely inferential character. This in turn negates the need to account for the practice of providing counter-examples. Any interlocutor attempting to provide a counter-example to a *logically* valid rule would, in general, be taken to have misunderstood the very nature of those expressions, specifically their purely inferential and hence universally valid character. Thus, the aspect of criticism by means of counter-examples poses no threat to (LCons) or (Val).

It seems, then, there is no reason to suppose that our proof-theoretic conception of logical consequences suffers from any conceptual mismatch, in the same vein that the argument from absolute semantics (AAS) refutes the LNLT, given how natural language semantics can account for linguistic understanding, and what model-theoretic conceptions require beyond this. As such, there is no conceptual mismatch to be feared for the inferentialist, hence no argument corresponding to the AAS could be looming in the background either.

Moving on to the extensional level, it seems then that any instance of somebody conjoining sentences with "and" from any set of premisses, and decomposing complex sentences linked with "and" into their constituent parts are all instances of a logical consequence relation. Of course, this is not to say that such instances are necessarily common. Consider the following exchange, taking place at a party, where the speakers must shout into each others' ears, despite standing next to each other:

A: "Tom, George and Sarah have arrived."

B: "Hey C, did you hear? George and Sarah have arrived."

C: "Cool, thanks for letting me know!"

D: "Hey C, has Sarah arrived yet?"

C: "Yes, she has."

Assuming we do not have a way to account for coordination effects, the inferences made by the interlocutors in the above example do not qualify as logically valid, since they are not instances of and-I or and-E. In other words, if we go with option of treating "and" only as a logical constant if it syntactically conforms to the initially presented rules, the circumstances in which speakers will explicitly structure their sentences in such a way will be rather uncommon. Nevertheless, the LNLT does not require logical consequences to arise commonly, but merely sometimes, and there can be no doubt instances of and-I as well as and-E can be found (cf. the 'Thomas-example' above).

Coming from these observations, we are now in a position to see what goes wrong with the ALC. From the inferentialist perspective, we simply lack anything analogous to (P3):

(P3) However, natural language semantics *are* absolute, as they could otherwise not account for linguistic understanding, which they should.

Recall that it was (P3) that squared badly with the requirements for logical notions (P1) and hence for natural language semantics (P2), in relation to the LNLT:

(P1) In order for any purported logical constant in natural language to generate a logical consequence relation, such a constant would have to pass the invariance test, hence would have to be identified as *logical* to begin with, *and* allow the tracing of truth-preservation in its associated entailments over a range of models.

(P2) In order for a logical constant to do so, natural language semantics would need to involve relativised truth-conditions, and hence could not be absolute.

Instead of having to pass any invariance tests, a purported logical constant's meaning-constitutive inference rules merely require a certain epistemic and formal profile in order to count as logical, and the corresponding entailments count as logical consequences just in case they are a chain of applications of such rules of inference. As I have argued above, none of these things take us beyond the realm of inferentialist (meta)semantics – be it accounting for linguistic understanding or providing the grounds and nature of meanings themselves. Thus, there is no surprise analogous to (P3) that would render the necessary conditions for logicality unsatisfiable with respect to the general set-up of inferentialist (meta)semantics. As such, we *can* find logical constants in natural language, which in turn *do* generate logical consequence relations – at least in principle, if not also on the extensional level.

This brings us to our main conclusion, namely that not only do proof-theoretic conceptions match on the conceptual level with general inferentialist assumptions concerning linguistic meaning, but there are, moreover, reasons to be optimistic about natural language *actually* containing logical consequence relations, specifically at least those involving meaning-constitutive inference rules of logical constants, of which we preliminarily identified at least one: "and". Thus, the LNLT is saved not only on a general conceptual level, but has been preliminarily vindicated on the extensional level as well.

6 Conclusion and Outlook

Let us summarise our findings. We started with the observation that the relationship between natural language and logic seems to be a close one, which made Glanzberg's arguments to the extent that there are neither logical consequence relations nor logical constants in natural language all the more puzzling. His primary observation was that in order for formal semantics to account for linguistic understanding, it cannot involve truth-conditions that are relativised to models. Rather, what we need are absolute semantics. However, model-theoretic conceptions of logical notions would require just such relativisations, hence natural language and logic come apart. We then briefly discussed Sagi's criticism of Glanzberg's arguments, and preliminarily found it wanting. It seems there is nothing apart from logical notions that could motivate a reintroduction of a sufficient number of models, but doing so begs the question against Glanzberg. Reaching a dead-end, we moved on to consider a different approach to linguistic meaning and logical notions. After giving reasons to be sceptical of the referentialist truth-conditional approach, we instead adopted an inferentialist stance both in metasemantics and semantics. In the former case, we took rules of inference to be the grounds on which expressions acquire their meaning, and also as the key ingredient in knowing the meaning of an expression. On the level of semantics, we then took a further step and equated meaning with the rules themselves. To get a handle on logical notions that accord well with our inferentialist framework, we further chose the corresponding proof-theoretic conceptions. These saw logical constants to be constituted by a set of inference rules with a specific epistemic and formal profile, and based on this, took logical consequence to be a sequence of such inferences.

With these tools in hand, we then set out to ascertain the logic in natural language thesis (LNLT) with respect to the package of inferentialist (meta)semantics and proof-theoretic conceptions. I argued that the proof-theoretic requirements for logical constants are compatible with the inferentialist's take on linguistic meaning, and further motivated the idea that we *can*, in fact, find logical constants in natural language, at least as far as the word “and” is concerned. On the basis of these considerations, I then argued that we can reconstruct logical consequence relations within natural language on the basis of such excavated constants. I pointed out that reasoning via logical constants both serves as a ‘rock-bottom’ for justifying inferences and plays an expressive role in making inferential patterns explicit, hence logic finds a natural place in our ordinary inferential practice. As such, we saw that the inferentialist and the corresponding proof-theoretic conceptions do not face a conceptual mismatch in the same vein as do their referentialist and model-theoretic cousins. Moreover, thanks to “and”, we can even be optimistic about the actual existence of consequence relations in natural language.

In order to arrive at this conclusion, I have made several choices along the way, related to the precise way of spelling out inferentialist ideas as well as proof-theoretic conceptions. Naturally, not all of these assumptions are universally shared, much less within the inferentialist camp itself. Furthermore, given the current status of proof-theoretic semantics, we were forced to stay on a conceptual level when discussing proof-theoretic conceptions. As such, it is entirely possible that certain details of these accounts *do* square badly with ordinary inferential practice, or an inferentialist understanding of linguistic meaning. Nevertheless, I hope to have shown that with the motivated package that I have provided, the LNLT can be saved. Perhaps this, coming from our observations concerning the methodology and epistemology of logic, serves as a further motivation to adopt precisely such a package – specifically one of inferentialist nature as opposed to a referentialist counterpart.

It should also be clear that this investigation is far from completed. First, our consequence relation in natural language is rather uninteresting. For this reason, it should be instructive to study further potential constants, such as “not”. If we were to find purely inferential, primitively obvious introduction and elimination rules for “not” as well, we might be in a position to reconstruct some form of propositional logic within natural language. From there, investigations into “or”, “if”-“then”, etc. could be launched. Once this would be completed, the next natural step would be to study quantification in natural language – a topic we lacked the space to delve into, unfortunately.

In addition to this, the epistemological and methodological observations from chapter 1 are deserving of further scrutiny. The match between inferentialism and proof-theoretic conceptions should allow for a reinvigoration of the idea that our knowledge of logic *is* knowledge of a specific part of our language. Thus, even though such knowledge would most plausibly count as *a priori*, it could be argued that it is nevertheless not founded on some mysterious capacity of rational insight, but rather on an appreciation of the inferential behaviour of certain expressions, perhaps subverting the usual rationalist-empiricist divide in the literature on the epistemology of logic.

Last but not least, the discovered match between inferentialism and proof-theoretic conceptions should justify the use of logical tools in studying linguistic meaning, at least as far as proof-theoretic semantics are concerned. Perhaps, this match might also serve to motivate the use of logical tools in mainstream formal semantics. If we can recover logical rules of inference, it might be that certain notions such as L-analyticity might stage a comeback, provided we can reconstruct them using inferential resources.

All these further developments must be postponed for another occasion, unfortunately. In the meantime, I hope to have made a contribution by demonstrating that not only does the relationship between natural language and formal logic seem to be close, but we *can*, at least in principle, recover the latter within the former.

7 Bibliography

- Aloni, M., Roelofsen, F. & Theiler, N. (2017): "What's Wrong with Believing Whether?", in: *Proceedings of SALT* 27, 248-265.
  Cited as "Aloni et al. (2017)".

- Beall, J., Restall, G. & Sagi, G. (2019): "Logical Consequence", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition).*
  URL = <https://plato.stanford.edu/archives/spr2019/entries/logical-consequence/>.

- Beaney, Michael (2021): "Analysis", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2021 Edition).*
  URL = <https://plato.stanford.edu/archives/spr2021/entries/analysis/>.

- Boghossian, Peter (2011): "Williamson on the A Priori and the Analytic", in: *Philosophy and Phenomenological Research* 82 (2), 488-497.

- Brandom, R. B. (1983): "Asserting", in: *Noûs* 17 (4), 637-650.

- Brandom, R. B. (1994): *Making It Explicit.* Harvard: Harvard University Press.

- Brandom, R. B. (2000): *Articulating Reasons.* New York/Cambridge (Mass.): Harvard University Press.

- Burgess, A. & Sherman, B. (2014): *Metasemantics. New Essays on the Foundations of Meaning*, Oxford: Oxford University Press.

- Cotnoir, Aaron (2018): "Logical Nihilism", in: N. Pederson, N. Kellen & J. Wyatt (eds.), *Pluralisms in Truth and Logic*, Basingstoke: Palgrave Macmillan, 301–329.

- Davidson, Donald (1967): "Truth and Meaning", in: *Synthese* 17 (3), 304-323.

- Dekker, Paul (2011): "Dynamic Semantics", in: C. Maienborn et al. (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 1, 923-945.

- Dummett, Michael (1991): *The Logical Basis of Metaphysics.* Cambridge: Harvard University Press.

- Dutilh Novaes, Catarina (2015): "A Dialogical, Multi-Agent Account of the Normativity of Logic", in: *Dialectica* 69 (4), 587-609.

- Dutilh Novaes, Catarina (2016): "Reductio ad absurdum from a dialogical perspective", in: *Philosophical Studies* 173, 2605-2628.

- Etchemendy, John (2008): "Reflections on Consequence", in: D. Patterson (ed.), *New Essay on Tarski and Philosophy*, Oxford University Press, 263-299.

- Gamut, L. T. F. (1991): *Logic, Language and Meaning. Volume 2: Intensional Logic and Logical Grammar*, Chicago/London: The Chicago University Press.

- Gentzen, Gerhard (1934): "Untersuchungen über das logische Schliessen", in: *Mathematische Zeitschrift* 39, 405-431.

- Glanzberg, Michael (2014): "Explanation and Partiality in Semantic Theory", in: A. Burgess & B. Sherman (eds.), *Metasemantics: New Essays on the Foundations of Meaning*.
URL = <https://www.semanticscholar.org/paper/Explanation-and-Partiality-in-Semantic-Theory-Glanzberg/7242ad8196c820f71e86eddea8e50f535a65a37c>

- Glanzberg, Michael (2015): "Logical Consequence and Natural Language", in: C. R. Caret & O. T. Hjortland (eds.), *Foundations of Logical Consequence*, Oxford University Press, 71-120.

- Glanzberg, Michael (2020, *forthcoming*): "Models, Model Theory and Modeling", in: G. Sagi & J. Woods (eds.), *The Semantic Conception of Logic: Essays on Consequence, Invariance and Meaning*, Cambridge University Press.
The page numbers refer to the penultimate draft available on:
URL = <https://michaelglanzberg.org/home/research/#inpress>

- Glock, Hans-Johann (2017): "Impure Conceptual Analysis", in: G. D'Oro & S. Overgaard (eds.), *The Cambridge Companion to Philosophical Methodology*, Cambridge University Press, 77-100.

- Glock, Hans-Johann (2018): "Why Rules Ought to Matter", in: Beran et al. (eds.), *From Rules to Meanings: New Essays on Inferentialism*, New York: Routledge, 63-80.

- Glock, Hans-Johann (2020): "Animal Consciousness", in: H. Appelqvist (ed.), *Wittgenstein and the Limits of Language*, Routledge, 197-222.

- Hanfling, Oswald (2000): *Philosophy and Ordinary Language. The Bent and Genius of Our Tongue*, London/New York: Routledge.

- Heim, I. & Kratzer, A. (1998): *Semantics in Generative Grammar.* Oxford: Blackwell Publishing.

- Hjortland, O. T. (2019): "What Counts as Evidence for a Logical Theory?", in: *Australasian Journal of Philosophy* 16 (7), 250-282.

- Hughes, G. E. & Cresswell, M. J. (1996). *A New Introduction to Modal Logic.* London/New York: Routledge.

- Künne, Wolfgang (2003): *Conceptions of Truth.* Oxford: Clarendon Press.

- Lepore, Ernie (1983): "What Model-Theoretic Semantics Cannot Do", in: *Synthese* 54, 167-187.

- Linnebo, Øystein (2004): "Plural Quantification Exposed", in: *Noûs* 37 (1), 71-92.

- Linnebo, Øystein (2017): "Plural Quantification", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2017 Edition).* URL = <https://plato.stanford.edu/archives/sum2017/entries/plural-quant/>.

- MacFarlane, John (2010): "Pragmatism and Inferentialism", in: B. Weiss & J. Wanderer (eds.), *Reading Brandom. On Making It Explicit*, Routledge, 81-95.

- MacFarlane, John (2017): "Logical Constants", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Winter 2017 Edition).* URL = <https://plato.stanford.edu/archives/win2017/entries/logical-constants/>.

- Miller, Alexander (2002): "What is the Manifestation Argument?", in: *Pacific Philosophy Quarterly* 83 (4), 352-383.

- Peacocke, Christopher (1987): "Understanding the logical constants: a realist's account", in: *Proceedings of the British Academy* 73, 153-200.

- Peregrin, Jaroslav (2014): *Inferentialism. Why Rules Matter*, New York: Palgrave MacMillan.

- Prawitz, Dag (1985): "Remarks on Some Approaches to the Concept of Logical Consequence", in: *Synthese* 62(2), 153-171.

- Prawitz, Dag (2005): "Logical Consequence: A Constructivist View", in: S. Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford University Press, 671-695.

- Prawitz, Dag (2006): "Meaning Approached Via Proofs", in: *Synthese* 148, 507-524.

- Prior, Arthur (1960): "The Runabout Inference-Ticket", in: *Analysis* 21 (2), 38-39.

- Rumfitt, Ian (2000): "'Yes' and 'No"', in: *Mind* 109 (436), 781-823.

- Rumfitt, Ian (2017): "Against Harmony", in: B. Hale, C. Wright & A. Miller (eds.), *A Companion to the Philosophy of Language, Second Edition* (Vol. 1), Wiley-Blackwell, 225-249.

- Russell, Gillian (2020): "Logic isn't normative", in: *Inquiry* 63 (3-4), 371-388.

- Ryle, Gilbert (1949): *The Concept of Mind.* Chicago: The University of Chicago Press.

- Ryle, Gilbert (1971): *Collected Essays. Volume II*, London: Hutchinson.

- Sagi, Gil (2020, *forthcoming*): "Considerations on Logical Consequence and Natural Language", *to appear in Dialectica.*

- Schroeder-Heister, Peter (2018): "Proof-Theoretic Semantics", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2018 Edition).* URL = <https://plato.stanford.edu/archives/spr2018/entries/proof-theoretic-semantics/>.

- Sher, Gila (1991): *The Bounds of Logic: A Generalized Viewpoint.* Cambridge, Massachusetts: MIT University Press.

- Shoesmith, D. J. & Smiley T. J. (1978): *Multiple-Conclusion Logic.* Cambridge: Cambridge University Press.

- Skorupski, John (2017): "Meaning, Use and Verification", in: B. Hale, C. Wright & A. Miller (eds.), *A Companion to the Philosophy of Language, Second Edition* (Vol. 1), Wiley-Blackwell, 73-106.

- Smullyan, R. M. (1995): *First-Order Logic.* New York: Dover Publications, Inc.

- Steinberger, Florian (2009): *Harmony and Logical Inferentialism.* PhD Thesis, University of Cambridge.

- Steinberger, Florian (2011): "What Harmony Could and Could Not Be", in: *Australasian Journal of Philosophy* 89 (4), 617-639.

- Steinberger, F. & Murzi, J. (2017): "Inferentialism", in: B. Hale, C. Wright & A. Miller (eds.), *A Companion to the Philosophy of Language, Second Edition* (Vol. 1), Wiley-Blackwell, 197-224.

- Stokhof, Martin (2007): "Hand or Hammer? On Formal and Natural Languages in Semantics", in: *Journal of Indian Philosophy* 35 (5-6), 597-626.

- Strawson, P. F. (1950): "On Referring", in: *Mind* 59 (235), 320-344.

- Strawson, P. F. (1952): *Introduction to Logical Theory.* London: Methuen & Co Ltd.

- Troelstra, A. S. & Schichtenberg, H. (2000): *Basic Proof Theory.* Cambridge: Cambridge University Press.

- van Dalen, Dirk (2013): *Logic and Structure.* London: Springer.

- van Eijck, J., Visser, A., Brasoveanu, A. & Nouwen, R. (2016): "Dynamic Semantics", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition).*
URL = <https://plato.stanford.edu/archives/win2016/entries/dynamic-semantics/>.
Cited as "van Eijck et al. (2016)".

- Williamson, T. & Rumfitt, I. (1999): "Logic and Existence" in: *Proceedings of the Aristotelian Society, Supplementary Volumes* 73, 151-203.
Cited as "Williamson (1999)".

- Wittgenstein, Ludwig (1921, 2013): *Tractatus Logico-Philosophicus.* Frankfurt am Main: Suhrkamp.
Cited as "TLP".

- Wittgenstein, Ludwig (1953, 2013): *Philosophische Untersuchungen.* Frankfurt am Main: Suhrkamp.
Cited as "PI".

- Yalcin Seth (2017): "Semantics as Model-Based Science", in: D. Ball & B. Rabern (eds.), *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics*, Oxford University Press, 334-360.

- Zimmermann, T. E. (1999): "Meaning postulates and the model-theoretic approach to natural language semantics", in: *Linguistics and Philosophy* 22 (5), 529-561.