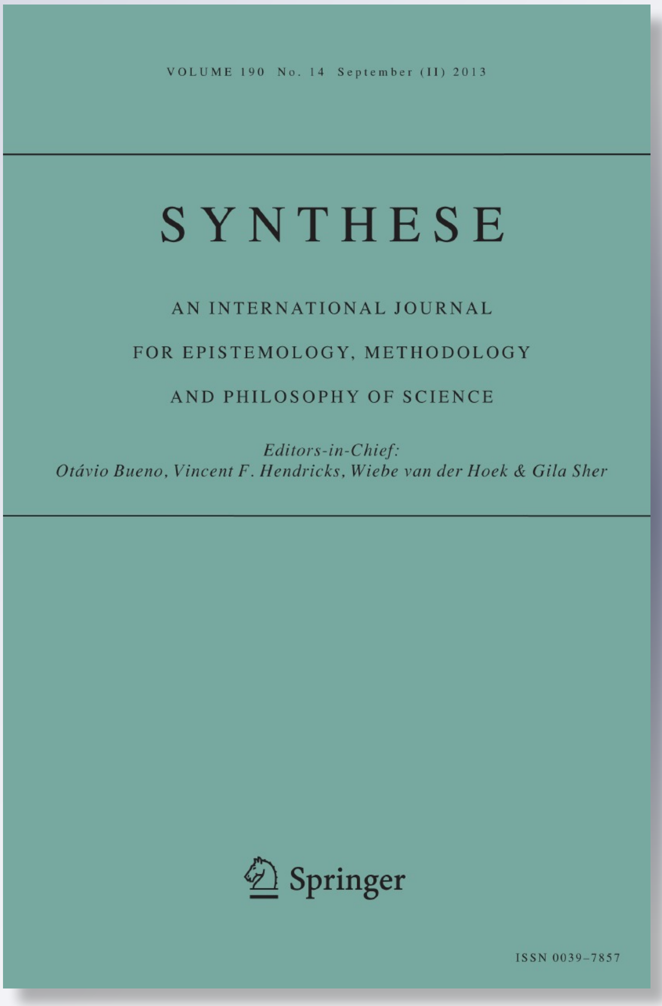
Criteria for logical formalization

Jaroslav Peregrin & Vladimír Svoboda

Synthese
An International Journal for
Epistemology, Methodology and
Philosophy of Science

ISSN 0039-7857
Volume 190
Number 14

Synthese (2013) 190:2897-2924
DOI 10.1007/s11229-012-0104-0



Springer

Synthese (2013) 190:2897–2924
DOI 10.1007/s11229-012-0104-0

Criteria for logical formalization

Jaroslav Peregrin · Vladimír Svoboda

Received: 29 September 2011 / Accepted: 26 March 2012 / Published online: 27 April 2012

Abstract The article addresses two closely related questions: What are the criteria of adequacy of logical formalization of natural language arguments, and what gives logic the authority to decide which arguments are good and which are bad? Our point of departure is the criticism of the conception of logical formalization put forth, in a recent paper, by M. Baumgartner and T. Lampert. We argue that their account of formalization as a kind of semantic analysis brings about more problems than it solves. We also argue that the criteria of adequate formalization need not be based on truth conditions associated with logical formulas; in our view, they are better based on structural (inferential) grounds. We then put forward our own version of the criteria. The upshot of the discussion that follows is that the quest for an adequate formalization in a suitable logical language is best conceived of as the search for a Goodmanian reflective equilibrium.

Keywords Logic · Logical form · Formalization · Reflective equilibrium

1 The nature of logical analysis

1.1 Logic and language

The relationship between the languages of formal logic and natural language is a complex and delicate one. Some of the formal languages, such as the language of

J. Peregrin · V. Svoboda (✉)
Department of Logic, Institute of Philosophy, Academy of Sciences of the Czech Republic, Jilská 1, 110 00 Prague, Czech Republic
e-mail: sv.oboda@seznam.cz

J. Peregrin
e-mail: jarda@peregrin.cz

classical predicate calculus or Montagovian intensional logic, are traditionally seen as useful tools for semantic analysis of natural language, while others, like the languages of some many-valued or modal logics, are quite detached from any natural means of communication. There is, however, a sense in which all logical languages worthy of the name must be anchored in the structures of natural language: formal systems whose languages lack this anchoring simply do not deserve, strictly speaking, the title *logical*. The point is that we can construct an abundance of formal systems that will employ "languages" furnished with something called "consequence" or "deduction system", or a kind of "semantics", but it is only to the extent that these artifacts reflect what we (in natural language) call *consequence, inference* or *semantics* that calling them thus becomes more than a deliberate fiat.[1]

Logic was conceived of to help us recognize and sustain the proper modes of argumentation, reasoning, justifying or carrying out proofs that inevitably take place in natural language (where *natural* language is not meant to be opposed to such devices as the language of ordinary mathematics, which is also to a large extent natural, but to *artificial* languages, constituted entirely by means of definitions). Natural language is, we claim, the principal area where meaning(s) as such get constituted[2] and therefore it must be the relationship to natural language that is *the measure of all logical things* (*of things which are logical, that they are, and of things which are not logical, that they are not*).

From this viewpoint, the assessment of the relationship between a given calculus and natural language is a crucial issue. We should have criteria for evaluating how fruitfully (if at all) we can use the calculus as a grid to be superimposed over natural language so as to explicate its inferential structure or semantics. And we should also have criteria for evaluating to what extent we can read the statements of the calculus as expressing the very kind of propositions that are expressed by sentences of our natural language.

Take modal calculi: Some of them, as numerous studies have shown, can be used to analyze the various versions of necessity and possibility that we have in mind when using the English words "necessarily" and "possibly" in various contexts. Interpreted formulas of the calculi can thus be seen as meaningful—as sentences treating necessities and possibilities in the very way English sentences do (perhaps in a less fuzzy manner). For other modal calculi no such relationship is in view. Their "statements" do not express what could be reasonably called propositions. Such calculi may be interesting from the mathematical viewpoint, but their status as *logical* systems—in the fundamental sense of the word—is questionable.

Hence, what is usually called *logical analysis* is not only a means of revealing or elucidating meanings of natural language sentences and assessing arguments formed out of them, it can also be seen as vindicating the status of the apparatus used for the analysis *as logic*. From this viewpoint, the criteria of assessing the proximity of

1 Of course, there is a broader construal of the term *logical* on the basis of which the term can encompass everything that is in any way related to logic (e.g. systems that arise out of some logical—in the narrow sense—system by means of purely technical modifications). But this is a secondary sense, which is marginal from our present perspective.

2 Adherents of other views of the nature of meaning, such as semantic Platonists, may want to say that it is a place where meanings become *available*, but this is not an issue that we want to delve into here.

a sentence of natural language and its alleged "logical structure" or "logical form" (viz. a formula of a formal language assigned to it in the process of logical analysis) are of basic importance. And it is surprising how little explicit attention the problem of criteria of an appropriate match between natural language expressions and their counterparts in logical languages receives in the (meta)logical literature. In fact, the only book-length explicit analysis of this problem we know of is that of Brun (2003).[3]

1.2 Atomism vs. reflective equilibrium

Recently, Brun's conclusions have been examined and reevaluated in an interesting paper by Baumgartner and Lampert (2008). The authors accept, in part, Brun's basic framework, but they argue for upgrading it substantially by moving it in a direction inspired by Wittgenstein's *Tractatus* (1922). In particular, they argue that the criteria of assessment of the match between expressions of natural language and expressions of formal language should be *semantic* (based on comparing truth-conditions), *atomistic* (the appropriateness of a logical analysis of a sentence should not depend on the analysis of other sentences[4]) and that all kinds of inferential dependencies among sentences of a natural language (be they exclusively a matter of logical vocabulary, or depend on extralogical parts of the sentences involved) should come out captured as *logical dependencies* within the target formal language. This is accomplished by means of a kind of "radical contextualism": how a sentence embedded in a given text is to be analyzed is determined by the particular vocabulary employed in the text.

One of the notions of logical analysis that Baumgartner and Lampert argue against is the notion known as the *reflective equilibrium*. Goodman (1955), who, as far as we know, first introduced this notion (though not the term) into this context, writes[5]:

> [D]eductive inferences are justified by their conformity to valid general rules, and […] general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend*. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either.

According to this picture, logical formalisms basically generalize and systematize the inferential and semantic features of natural language and so are liable to similar kind of criticism as other empirical generalizations. However, due to the fact that natural language is vague, open-ended and not necessarily entirely consistent, formalization

3 Relevant considerations can also be found in Blau (1977, 2008), Sainsbury (1991), Epstein (2001) and Svoboda and Peregrin (2009).

4 Some of the formulations of Baumgartner and Lampert indicate that what they consider as an "atom" is not a sentence, but a "text". However, what they consider a "text" is always expressible as a truth-functional composition of sentences (see below), and hence obviously as a complex sentence. Moreover, their criticism of holism seems to indicate that they would not settle for anything short of atomism in the traditional sense.

5 See, e.g., Resnik (1985) for a more detailed exposition of logic as based on the reflective equilibrium.

also does the job of sharpening, explicating and removing inconsistencies; as a result it gains a certain normative authority over our use of natural language (at least in contexts requiring accuracy of articulation). Just like we say that to call a fish-like animal with lungs *fish* is "improper" because we concluded, by means of generalization (and a systematization of generalizations into a classification), that fish have gills, we can say that a particular inference, particular argument, or particular proof is incorrect because we concluded, by means of generalization and systematization, that this is not how proper inferences generally go.[6]

In this paper we want to point out that, though both the discussion given by aumgartner and Lampert and the framework they arrive at are in many respects novel and interesting, their proposal creates more problems than it solves. We will argue that the atomism they promote is a mere illusion and that once we see that a kind of holism is inevitable, we are back with the reflective equilibrium picture. But this picture, as we found it in the literature, is essentially underdeveloped—as far as we know, it has never been elaborated in detail. In this paper, we will try to flesh it out.

The structure of the rest of the paper is as follows: in Sect. 2 we give an overview of Baumgartner and Lampert's "*new picture* of adequate formalization"; subsequently, in Sect. 3, we point out its problematic features. In Sects. 4 and 5 we present our own account of the process of logical formalization and of the criteria that govern it. We defend a version of the account rejected by Baumgartner and Lampert—we view logical formalization as a matter of homing in on a reflective equilibrium. In Sect. 6 we summarize our views and compare them with those of Baumgartner and Lampert. Section 7 contains the paper's conclusion.

2 Baumgartner and Lampert's "new picture"

2.1 Logic as *ars iudicandi* vs. as *ars explicandi*

Once we admit that any formal system deserving the name *logic* must be, in some way, anchored in natural language, the question of how formulas (statements of the logical language) relate to natural language sentences turns out to be central. This question can be split into two complementary questions—the question of the best "translation" of natural language sentences (or arguments) into a logical language, and the question of the best converse translation of logical formulas into a natural language. The first question is, of course, primary and much more common. When we ask what is the best rendering of a given sentence, argument or text of natural language (let say English) in a logical language, we face the question of its adequate *formalization*.[7]

[6] This provides for a *normativity* of logic, for it gives logic the power of telling us, in some cases, what are the correct inferences in natural language and what are not. Notice, however, that there is at least one more sense in which logic is normative: unlike in the case of fish, the classification concerns *rules*, i.e. normative entities.

[7] Here we will, for the most part, view the problem of an adequate formalization as the problem of the best formal representation of a sentence, argument, or text in the language of a given logical calculus. But we can also approach it more generally and search for the most adequate formalization across available logical languages (we will return to this later).

Any student who has gotten a passing grade in their elementary logic class knows that some formalizations of a given sentence are classified as correct, and others as incorrect. The criteria are usually not made fully explicit during the classes, but it is normally supposed that the teacher is unproblematically capable of determining the correct formalization. (Thus, it is taken for granted that when a pupil is asked to provide a logical analysis of, say, *A book is on a table*, she will come up with something like $\exists x \exists y (B(x) \wedge T(y) \wedge \mathrm{O}(x,y).)$)[8] Though the practice of formalization is quite common, only a few attempts have been made to present explicit criteria of this kind of correctness.[9]

Baumgartner and Lampert have developed their *new picture* of adequate formalization as their original theory of this. Their main point is to demonstrate that "logic is not to be seen as *ars iudicandi* capable of evaluating the validity or invalidity of informal arguments, but as *ars explicandi* that renders transparent the formal structure of informal reasoning" (p. 93).[10] Although we are not unsympathetic with this idea (we too think that logic must be based on "rendering transparent the formal structure of informal reasoning"), we find the picture that they put forward problematic in several respects. Before we articulate our objections and offer an alternative picture we will try to summarize their theory.

2.2 The syntactic criteria

Baumgartner and Lampert say that the process of formalization of a text[11] must be preceded by a decomposition of the text into a complex of statements combined together in a truth-functional way. They call this step *interpretation* and stress that it is not supposed to be based only on a syntactic analysis but also presupposes the ordinary understanding of a competent speaker.[12] The crucial step of formalization itself, then, consists in translating the individual statements (sentences that are either true or false) into a logical language.

The underlying idea of the approach of Baumgartner and Lampert is that "logical formalization cannot be separated from semantic analysis" (p. 95). Thus, formalization is a process in which insights concerning the form of the expressions and their content are intertwined. The correctness of formalization can be assessed using *syntactic* or *semantic* criteria. Though Baumgartner and Lampert choose the semantic criterion as the one that is more satisfactory, let us first briefly review also the criterion they call syntactic. They proceed by formulating two syntactic subcriteria—*correctness* (in a narrower understanding) and *completeness* (pp. 96 and 103):

[8] Let us leave aside, for the moment, the extent to which the teachers *follow* some criteria and the extent to which they *make* them.

[9] Aside from Brun's book cited above see, e.g. the classical exposition due to Sainsbury (1991).

[10] Hereafter, all page numbers refer to Baumgartner and Lampert (2008).

[11] The simplest text that can be formalized is usually a sentence but, in principle, we could also speak about the formalizing of phrases or even individual words. Normally, however, texts consist of several sentences.

[12] How exactly is this kind of formalization supposed to be accomplished in the case of different texts is not entirely clear.

(COR) The formalization Φ of a statement A is correct iff every inference S, such that the formalization Ψ of S contains Φ as a premise or a conclusion and Ψ is formally valid, is informally valid.

(COM) The formalization Φ of a statement A is complete iff every formalization Ψ, such that Ψ is the formalization of an inference S and contains Φ as a premise or a conclusion and S is informally formally[13] valid, is formally valid.

While (COR) is, according to Baumgartner and Lampert, quite standard, (COM) is not commonly seen as a requirement that adequate formalizations have to meet. The two main reasons why completeness is omitted are (a) the termination problem (an unlimited number of inferences are in play), and (b) the problem that the aspiration to completeness blurs the distinction between logical (formal) analysis and semantic analysis.

The first problem Baumgartner and Lampert solve (or more precisely reduce) by a modification of (COM) assuring that the required parallel inferences are limited only to inferences concerning terms explicitly mentioned in the realizations of the involved formalizations (p. 106).[14]

(COM′) The correct formalizations $\Phi_1, \ldots, \Phi_n$ of $A_1, \ldots, A_n$ are complete iff every formalization Ψ, such that Ψ is the formalization of an inference S, that does not consist of premises or conclusions other than $A_1, \ldots, A_n$ or their negations or of verbalizations of formulae exclusively composed of the categorematic expressions mentioned in the realizations of $\Phi_1, \ldots, \Phi_n$, and S is informally valid, is formally valid.

The second problem they face head on. They are convinced that the distinction between those inferences that are valid due to their logical form (i.e. are valid logically), and those that are valid due to meaning of the extralogical expressions they contain, is dubious and so a proper formalization should (un)cover *both* kinds of inferences.[15] The completeness of adequate formalizations, in their understanding, consists in the fact that non-formal (material) dependencies among statements are completely mirrored by (logical, classically understood) formal dependencies.

With recourse to (COR) and (COM′), correctness and completeness of the formalization of a text T composed of statements $A_1, \ldots, A_n$ is syntactically defined as follows (p. 107):

(SYN_T) The formalization Φ of a text T is correct and complete iff (COR) and (COM′) are satisfied for all formalizations $\Phi_1, \ldots, \Phi_n$ of $A_1, \ldots, A_n$ and (COR) is satisfied for Φ of T, such that Φ is a truth-function of $\Phi_1, \ldots, \Phi_n$.

[13] Aside from formal validity (viz. defined validity of statements of a formal system) and informal validity (intuitive validity of sentences of natural language), the authors introduce the concept of "informal formal" validity (informal validity due to form only).

[14] Realizations are a matter of associations of non-logical symbols appearing in formulas with meaningful terms. More about this aspect of their theory later.

[15] They associate this idea with Wittgenstein's *Tractatus* (1922) and assume that expressions assigned to the categorematic parts of any formula are always *logically independent* (p. 107). They present this limitation of the (realizations of the) formal language as a consequence of (COM′), but there does not seem to be a clear connection.

2.3 The semantic criteria

Baumgartner and Lampert suggest that this syntactic criterion points in the right direction but, as the application of (COR) and (COM′) does not terminate and involves circularities, they prefer a semantic criterion. Its definitive version is concisely presented by the following formulation (p. 112):

(ADS) The formalization Φ of a text T is adequate iff (TC'_T) is satisfied and Φ is at least as similar to T as any other equivalent formula that satisfies (TC'_T).[16]

The criterion (TC'_T) referred to by (ADS) says (p. 109):

(TC'_T) The formalization Φ of a text T is correct and complete iff (TC′) is satisfied for all formalizations $\Phi_1, \ldots, \Phi_n$ of $A_1, \ldots, A_n$ and for Φ of T, such that Φ is a truth-function of $\Phi_1, \ldots, \Phi_n$.

And the criterion (TC′) referred to by (TC'_T) in turn says (p. 108):

(TC′) The formalization Φ of a statement A is correct and complete iff relative to all interpretations $\Im$ of Φ, Φ has the same conditions of truth and falsehood as A has according to the informal judgement.[17]

The criterion that Baumgartner and Lampert propose is *atomistic*, which means that the adequacy of a formalization of a sentence does not depend on formalization of other sentences. And, indeed, this is what Baumgartner and Lampert consider its basic *virtue* for, as they claim, holistic criteria lead to the situation where "every seemingly correct formalization constantly risks to be identified as incorrect in light of other formalizations" and hence the criteria cease to be true *criteria* and become mere "rules of thumb" (p. 97).

3 The pitfalls of the "new picture"

3.1 The ambiguity of "formalization"

Let us now consider Baumagarter–Lampert's approach in more detail. We can divide our criticism into several points. In this subsection we are going to point out a significant (and perhaps symptomatic) ambiguousness affecting their proposal, which tends to obscure their exposition in its entirety. In the following section we will concentrate on the problems brought about by their radical contextualism and the Tractarian grounding of their approach. In Sects. 3.3 and 3.4 we will cast doubt on the idea

[16] Similarity is measured by identifying syntactic counterparts: A syntactic feature of a formula must have a counterpart in the statement and a syntactic feature of a statement must have a counterpart in the formula.

[17] Let us note that there is apparently no distinction between a statement and a text within Baumgartner and Lampert's approach, as (TC′), which concerns a formalization of *statements*, is within (TC'_T) also applied to a formalization of a text T. (In fact it is unclear why the requirement that (TC′) is satisfied for all the particular statements constituting T is mentioned separately in (TC'_T)—if (TC′) can be applied on the whole text, then (TC'_T) seems superfluous.)

that semantic criteria of correct formalization are to be preferred to the inferential ("syntactic") ones.[18]

We can begin with a simple English sentence and its formalization that Baumgartner and Lampert mention as Brun's example of formalization (p. 98):

(1) *All animals have heads.*

(1′) $\forall x(F(x) \rightarrow G(x))$

F : ... is an animal; G : ... has a head

The second line of (1′) is what Baumgartner and Lampert call *realization*. A device of this kind was introduced by most of the authors considering the problem of logical formalization[19]: it links non-logical symbols of the formula resulting from the formalization to the expression of the natural language it supersedes. If we take formalization as a kind of translation from natural language to a formalized language, then the *raison d'etre* of this device is clear: to formalize (1) as $\forall x(F(x) \rightarrow G(x))$ with F standing for *is an animal* and G standing for *has a head* appears to be reasonable, while to formalize it as the same formula, but with F standing for *has a head* and G standing for *is an animal* seems weird. But Baumgartner and Lampert put this device to an additional use: for them it opens the way to an oscillation between seeing the outcome of formalization as an uninterpreted formula (with F and G being merely parameters) and as an interpreted one (with F and G being interpreted via the correspondence schema), and this is exactly what Baumgartner and Lampert, in effect, do. They, on the one hand, distinguish between *formalization* (the formula itself) and its *realization* (the annotation at the second line),[20] but subsequently they say that it is the two lines that constitute a formalization (see p. 101). Later on they add that formalizing a conclusion amounts to providing a formula representing the conclusion, which (as the realization is not part of the formula) seems to suggest that only the formula itself counts as formalization.

This ambiguity is, we are afraid, pernicious, especially when we come to Baumgartner and Lampert's methodology. If we focus—to make things more tractable—just on the formalization of a single sentence and take (1) as the example, we can apply the central criterion (TC′) to the particular case in the following way:

(TC$^{(1)}$) The formalization $\forall x(F(x) \rightarrow G(x))$ of the statement *All animals have heads* is correct and complete iff relative to all interpretations $\mathfrak{I}$ of $\forall x(F(x) \rightarrow G(x))$, this formula has the same conditions of truth and falsehood as the sentence *All animals have heads* has according to the informal judgement.

This sounds quite strange: How could the truth conditions of a natural language sentence co-vary with interpretations of the formula $\forall x(F(x) \rightarrow G(x))$? The sentence

[18] As we consider the term "syntactic criteria" potentially misleading, we will use the term "inferential criteria".

[19] Blau (1977, p. 7) calls it "assignment" ("Zuordnung"), Sainsbury (1991, p. 51) and Brun (2003, §6.1) call it "correspondence scheme".

[20] Realization assigns expressions of natural language to the categorematic parts of a formula (p. 97; see also Epstein (2001, p. 13), from whom Baumgartner and Lampert take over the concept). Baumgartner and Lampert require that expressions assigned to the categorematic parts of a formula by its realization are *logically independent* and neither tautologous nor contradictory (p. 107).

apparently does not contain any part which an interpretation would affect. And, indeed, how could the formula have truth conditions relative to an interpretation? It would seem that once it is interpreted (by the realization) is has a truth *value*, not truth *conditions*. What we learn from Baumgartner and Lampert is that the truth conditions to be evaluated in case of natural language statements are "generated by paraphrasing the formal interpretations with recourse to the realizations of the corresponding formulae" (p. 109). It is, however, unclear what this means.

It seems that the only thing that Baumgartner and Lampert can sensibly have in mind while promoting (TC′) is not the variation of interpretations in the first-order sense, but rather just the variation of extensions depending on circumstances. Perhaps what we should imagine under a realization of a predicate parameter like F is not an assignment of a set of individuals, but an assignment of a function mapping possible worlds on sets of individuals, as in an intensional logic? Under this—charitable?—reading, (TC$^{(1)}$) says that the formula and the sentence have the same truth value in all circumstances and hence the same conditions of truth and falsehood in all possible worlds (i.e. they denote the same proposition). But if this is the correct construal of (TC$^{(1)}$), it is somewhat unclear why (TC′) should be formulated in such an oblique way and why it mentions interpretations at all. It would seem that Baumgartner and Lampert could have used a much more perspicuous formulation like:

(TC*) A formalization Φ of a statement A is correct iff Φ and A have the same conditions of truth and falsehood.[21]

Criteria of this kind seem, at first glance, comprehensible and reasonable. But, under a closer scrutiny, we will notice that we still face serious problems if we adopt them. We will return to these problems later.

Reflecting on the ambiguity of (1′), we can conclude that the language in which Baumgartner and Lampert want to articulate the results of formalization (logical analysis) vacillates on the borderline between a language that is formal in the sense of having uninterpreted parts (parameters), and a fully interpreted language. Moreover, the way their semantic criterion is formulated is untenable. Ambiguities of this kind are, we are afraid, confusing and potentially misleading, to the point of devaluation of the whole process of logical analysis based on the theoretical apparatus.

3.2 Contextualism

Now we will turn our attention to the radical kind of contextualism which is typical of Baumgartner and Lampert's new picture of formalization and to the Tractarian grounding of the conception. The authors put forward the view that "whatever is informally valid must be shown to be valid on formal grounds by means of logical formalization" (p. 105). Within the atomistic framework that they accept, this ambition does not lead them to search for a formalization that would allow the uncovering of all informal inferential links of a given sentence to other sentences of its language (which is clearly a task that cannot be generally completed), but it results in the view that formalization is adequate if it allows for the treatment of all the *relevant* informal inferences

[21] Which is, in fact a concise formulation of their criterion that they mention on p. 98.

as logically valid. Which informal inferences are relevant depends on the conceptual (terminological) buildup of the particular sentence or text that is being formalized. Thus, within the context of different sentences, the same sentence may be adequately formalized by very different formulas.

So if we face the problem of formalizing the sentence

(2) *Cameron is a mother.*

within the framework of classical predicate logic (CPL) according to Baumgartner and Lampert's "new picture of adequate formalization" we easily fall into trouble.

Presumably, the formalization of (2), standing alone, would be

(2′) $F(a)$
F : ... is a mother; a : ... Cameron

Let us now assume that we face the problem of formalizing the text:

(3) *Cameron is a mother. Therefore, Cameron is a woman.*

If we subject the sequence of sentences to interpretation[22] we will get, according to Baumgartner and Lampert's suggestion:

(3*) *Cameron is a mother* $\rightarrow$ *Cameron is a woman*[23]

which can then be formalized as

(3′) $G(a) \wedge H(a) \rightarrow G(a)$
G : ... is a woman; H : ... has a child; a ... Cameron[24]

Baumgartner and Lampert claim that this formalization is complete and accept it as adequate. Their analysis, however, raises questions. First it is clear that (3) informally entails (2). However, (3′) does not entail (2′).[25]

Things get even more problematic if we consider the text:

(4) *Cameron is a mother. Bob is a widower.*

If we subject the sequence of sentences to interpretation we get

(4*) *Cameron is a mother* $\wedge$ *Bob is a widower.*

But it is clear that (4*) cannot be formalized as

(4′) $F(a) \wedge K(b)$
F : ... is mother; a : ... Cameron; K : ... is widower; b : ... Bob

---

22 Here the term is used in the specific sense introduced on pp. 96–97 of Baumgartner and Lampert's article.

23 An obvious objection is that concatenating the two sentences by material implication distorts the meaning of (3), as it does not capture the character of the relation between the two sentences. We will see that this paraphrasing is indeed hardly justifiable.

24 For this proposal of formalization of the argument see p. 101. Unlike Baumgartner and Lampert we use brackets around individual parameters.

25 Moreover, even if we waive the principal inadequacy caused by using an implication for capturing the "therefore", the proposed decomposing of the meaning of term "mother" is also far from unproblematic. For example, if it were correct, the sentence *Lassie is the mother of two nice puppies* would have to be analytically false.

This attempt does not comply with the (IN) requirement that is postulated on p. 107 of the article. The requirement demands that the expressions assigned to the categorematic parts of a formula by its realization are *logically independent*. The predicates *is a mother* and *is a widower* are clearly *not* independent—somebody who is a mother cannot be a widower and vice versa. But the predicates *is a woman,. has a child, is a man, his wife died* do not fare any better. In fact, it is entirely unclear what kind of formula would be suitable in this simple case and, hence, what formalization of the sentence *Cameron is a mother* would be adequate if it appears as part of (4).

In general, we can say that it is very difficult to find any truly atomic sentences or predicates in natural language, hence the task of identifying them for even a limited context might well be a mission impossible. Not much less problematic is the fact that any adequate analysis of a given sentence may suddenly turn out to be inadequate when we find out that the sentence is, in fact, followed by another one. We don't think that contextualism of this kind is acceptable.[26]

3.3 Semantics as foundation

The problems we have discussed so far are specific for Baumgartner and Lampert's approach; it seems quite likely that they could be avoided without relinquishing the basic intuitions connected with the semantic criterion. After all, (TC*) appears to be a rather natural point of departure of the formalization project and several scholars theorizing about it have put forward criteria based on the same fundamental idea.[27] We are, however, convinced that giving priority to the semantic version of the criterion of adequate formalization is a strategy that introduces more problems than it solves. To indicate them we will now critically examine reasons that might substantiate the conviction that semantic articulations of the criterion of adequate of formalization are preferable to the inferential ones.

It is hardly controversial to say that it can be useful to consider the adequacy of logical analysis in semantic terms. It is clear, for example, that if we represent possible situations by means of truth tables we get a handy method of elimination of some unacceptable formalizations. For example, if we are asked to formalize the sentence *It is rainy and cold outside* we can easily identify the situations (rows in the truth table) in which it is true and demonstrate that trying to formalize the sentence by formula $A \vee B$ ($A \ldots$ *It is rainy outside*, $B \ldots$ *It is cold outside*) would lead to a wrong analysis—the sentence simply is not true in the same situations as the formula.

But we should keep in mind that even if we stay on the level of propositional logic, the semantic considerations of this kind are only of a limited use. Let us suppose that we want to identify the correct formalization of the sentence *If it is rainy, then it is wet* in the language of classical sentential calculus. It is obvious that we are likely to get

26 A contextualism that we would be ready to concede is a different one: We are convinced that the only context that may influence the adequacy of formalization is its purpose. If we analyze a text, e.g. for the purposes of a computer implementation of its semantics, the desired output may be very different than if our goal is just to show the untenability of an argument to an opponent.

27 Epstein (2001) and Sainsbury (1991) consider the adequacy criteria more or less only on the semantic level; whereas Brun (2003) considers both semantic and inferential criteria.

widely varying answers if we ask competent speakers in which of the situations represented by the rows of the table this sentence is true, and in which it is false. Some may claim that the sentence is truth-evaluable only in situations when it rains and hence is true only when it rains and it is wet; others would claim that though the sentences might be perhaps judged as false if it rains and it is not wet, any other situation alone does not determine its truth value uniquely; and yet others would say that it is true in all these other situations, i.e. situations in which it does not rain or it is wet. (Speakers of the last category would likely be diligent graduates of logic classes.)

Hence, if we insist on sameness of truth conditions, we have to conclude that an adequate formalization of the sentence in the language of the classical sentential calculus is impossible. Most logicians, however, would, without much hesitation, propose the formalization $A \rightarrow B$. This does not mean that they really do think that the truth conditions of the formula are the same as those of the sentence. They only take it for granted that it is the best option available. And if they were asked why they go for this formalization, they would probably not talk about truth conditions at all (they might, instead, refer to the fact that such a formalization has turned to 'work' for practical purposes). This indicates that, though what adherents of (TC*) should do is seek a formula (or, more precisely, a formula made truth-apt by means of something similar to the correspondence scheme) with truth conditions matching those of the sentence to be analyzed, what they *really* do is adapt their understanding of the truth conditions of the sentence and the formula so that they match each other. (Otherwise, they would be forced to say that most sentences of natural language cannot be adequately formalized in languages of common logical systems.)

This fact indicates that what is in question is not just *comparing* the two independent sets of truth conditions, but also giving consideration to their *adaptation to each other*.[28] This is, of course, not an objection that would disqualify the semantic criterion as compared with its alternatives—no other criterion could do better in this respect, because logical analysis is not a process of reflecting ready-made structures of natural language, but rather a process that also tampers with the structures. However, it casts essential doubts on the kind of 'clean' picture of the semantic criterion that Baumgartner and Lampert (and perhaps others) fancy.

However, there are also problems peculiar to the semantic formulation of the criterion. The first—and most obvious—problem is that the semantic criterion cannot serve as a good guideline for formalization of necessarily true sentences: any tautology of a formal language comes out as a suitable formalization of any sentence that is analytically true according to (TC*).[29]

More problems arise when we move to formalizations in terms of the language of the predicate calculus, the semantic criterion (TC*) faces deeper (and more idiosyncratic)

[28] This is noted by Epstein (2001), who formulates his analogue of our (TC*) as "*we can convince ourselves* that the formalization and the original proposition would both be true or both false relative to any given universe" (p. 16, our emphasis) because "for informal propositions what we *mean* by 'truth-conditions' is (or should be construed as) exactly what we get in the rewriting part of formalizing" (p. 167).

[29] Someone could argue that the same problem arises for inferential criteria—if the inferability relation is construed simply as truth-preservation, then any analytically true sentence is inferable from any sentence. But this only shows that inferential criteria must be based on a stronger conception of inferability. It is intuitively clear that "All bachelors are men" is not inferable from "$7 + 5 = 12$".

problems. How could we employ this criterion to decide whether $\forall x(F(x) \rightarrow G(x))$ is a/the correct formalization of *All animals have heads*? It may seem, at least prima facie, that the criterion asks us to examine each possible world (circumstance) and find out whether the formula (equipped with the corresponding realization) and the sentence have the same truth value within it.

Of course, we cannot do this literally—we cannot check the infinite number of cases world by world. Already the idea of trying to identify, in every possible world, the set of animals, that of individuals equipped with a head, and then look at their relationship, looks foolish. If we deal with logical problems we are, naturally, not expected to inspect the world (or its alternatives) and collect empirical data (and even if we were, we could hardly be expected to do it for an infinity of cases). Checking the adequacy of logical analysis is not a matter of checking states of the world—what is substantial is that we have an (objective) criterion that determines whether a particular formalization Φ of a particular sentence A is adequate—the only 'empirical' data we might need to know concern the exact meaning of A.

The idea, of course, is that we do not consider individual worlds—we, in fact, merely compare some characterizations of the classes of worlds in which the sentence in question is true. (As Baumgartner and Lampert spell it out, we only compare *structural descriptions* of the respective classes of worlds or circumstances in which the sentence resp. the formula are true. In the case of (1) resp. (1′), this would amount to checking whether the descriptions "the worlds in which the set assigned to F is part of the set assigned to G" and "the worlds in which all animals have heads" pick up the same class of worlds).

This approach leads us to the problem of delimiting "suitable interpretations" of formulas. Thus, for example, if we assume that (1) is analytically true (let us do so for the purposes of this discussion), we have to exclude every interpretation of (1′) that would not map F on a subset of G.[30] (If such interpretations were not excluded, then there would be possible worlds in which (1′) would be false.) Such an exclusion, however, brings us beyond the boundaries of logic, for they stem from the relationship of the extralogical terms *animal* and *head*. (Baumgartner and Lampert try to turn it into a logical matter in the well-known Tractarian way.[31])

3.4 Semantics vs. "syntax"

However, we think that this approach to comparing truth conditions faces a problem deeper than that of *how* to exclude "unsuitable interpretations". The deeper problem is *why* to exclude just the interpretations we exclude, i.e. why consider some interpretations "suitable" and other "unsuitable". An answer seems to be at hand: We need to exclude, and hence render "unsuitable", those and only those interpretations which

[30] The sentence, however, might be also taken as analytically false since, for example, corals and sponges count—according to a well established classification—as animals too.

[31] The idea is that if we manage to descend, by means of our analysis, to the level of genuine atoms, no such exclusion will be necessary. (If we, for example, find out that $F(x)$ in (1′) can in fact be decomposed into $G(x) \wedge J(x)$, then any interpretation which would not map F on a subset of G is excluded purely formally).

would spoil the equivalence of the truth conditions of the formula with the formalized sentence. But, in such a case, the "semantic criterion" does not seem to be any criterion at all and becomes a pure stipulation. Hence, if above we concluded that in applying the semantic criteria we adapt our understanding of the truth conditions of the analyzed sentence to match the analyzing formula,[32] here we see that something similar happens in the opposite direction: The truth conditions of the formula are carved, by means of a gerrymandering of the set of "suitable interpretations", so as to match those of the sentence.

The question, then, is to what extent can we really *check* the coincidence of the truth conditions and to what extent can we *make* them coincide. Hence, this leads to a more general question of whether we can have a semantic *criterion* that would be essentially superior to the inferential ("syntactic") one (in the sense that the former would grant us something beyond the reach of the latter), provided that all the problems just hinted at could be waived (be it dispensing with the termination problem, as Baumgartner and Lampert suggest, or something deeper, as is sometimes claimed). As a thorough discussion of this topic would require another article, we present it here in an abbreviated form.

To show that the semantic criterion is superior to the inferential one would amount to demonstrating that the former can help us achieve something we cannot achieve with the latter, i.e. that it can help us show that some features of the formulas which are candidates for the role of an adequate formalization of a sentence A (let say *All animals have heads*) cannot be identified by inferential means, whereas they can be captured on the semantic level. But can such a situation really arise? To consider this let us assume that we have two candidate formalizations of A, Φ and Ψ, which are inferentially indistinguishable but distinct with respect to their semantic features.

Assuming that the two formulas are semantically distinct though, at the same time, inferentially indistinguishable, we assume that there is a semantic distinction that does not influence inferability of the formula from other formulas and inferability of other formulas from the formula. How could this arise? There seem to be two possibilities. One is that there *are* some differences between the inferential roles of Φ and Ψ, but they are not manifest in the language under consideration, but only in a metalanguage. Another is that there are no differences in inferential roles whatsoever (not in the object language, nor in any kind of its metalanguage), but still differences in truth conditions. In the second case, we must obviously assume that the circumstances in which the difference between the respective truth conditions of Φ and Ψ become actual are not characterizable in language (for otherwise we would have a sentence or a set of sentences from which only one of Φ and Ψ would be inferable) and, hence, that the semantic difference between Φ and Ψ is a matter of pure, practical know-how. This case, then, cannot be helpful if what we are after is a *theory*, and hence cannot be used to justify the semantic criterion.

Thus, defenders of the superiority of the semantic criterion are left with a case where the semantic difference does have an inferential manifestation, though merely

[32] We can also quite well imagine that an adherent of (TC*) would be ready to stipulatively neglect the significant difference between meanings of the sentence *If it is rainy, then it is wet* and of the (realized) formula $A \rightarrow B$, and adopt the view that they share their truth conditions.

in a metalanguage.[33] Now, however, remember that Φ and Ψ are mere artifacts, which were created together with a framework of a formal language (e.g. first-order predicate calculus). So any 'meaning' they may have is a product of the stipulative extension of the calculus. This is to say that the only difference on the level of the metalanguage may be a stipulation of the kind of "let Φ be true in circumstances c^{Φ}" and "let Ψ be true in circumstances c^{Ψ}". Now why would one want to stipulate something like this? Perhaps to make a formula an adequate formalization of a sentence; for example, by stipulating "let Φ be true in circumstances when all animals have head"? But then that which the semantic criterion would bring us over and above the inferential one would be utterly trivial; it would not be significantly more interesting than the stipulation "let Φ be the adequate analysis of *All animals have heads*". In other words, we are back with the situation discussed in the previous section, where the adequacy of logical analysis is an uninteresting (for reason of being a purely stipulative) matter.[34]

Thus we don't see any reason why the semantic criteria of correct formalization should be superior to the inferential ones. In fact, in the rest of this paper we will suggest that inferential criteria are preferable. One of their main advantages is that they (unlike the semantic criteria) can provide a humanly manageable method of comparing acceptability of different formalizations. We should not forget that the problem of identifying a suitable (adequate) formalization is a *practical* problem. Thus a conception that offers practical guidelines for assessment of acceptability of different formalizations is more valuable than a conception offering a complex and fancy sounding criterion that is not associated with any humanly manageable (nontrivial) method of its practical application.

4 What is it that we do when we do logical formalization?

4.1 Regimentation

Before we turn to the outline of an alternative to Baumgartner's and Lampert's approach, let us make a general overview of what is in play when we speak about formalization. We will try in the following paragraphs to give an account of the intermediary steps that characterize usual formalization practices. At the same time we will introduce a terminology that we will need in the subsequent sections.

Let us assume that we have some basic stock of logical constants resulting from what we will call *rectification*, viz. replacing suitable natural language expressions by artificial signs with the aim of making inferential functioning of the original expressions definite and more transparent.[35] We assume that the constants have some kind of

33 Brun (2003, p. 210) notes that the criterion "can be applied only where the metalanguage in which we argue for or against a formalization of a statement of ordinary language possesses a suitable conceptual apparatus."

34 Epstein (2001, p. 166) mentions an extreme case of such a stipulative doing justice to the semantic criterion—any complex sentence can be formalized as atomic.

35 Thus, for example, rectification of the English connective *and* may lead us to the well-known symbol $\wedge$ with an exactly defined function fixed either by associating $\wedge$ with the well-known truth table, or as a result of positing certain axioms or rules. Employment of the symbols is typically associated with a suitable

semantics—be it explicit semantics, like the one based on truth-tables, or implicit ones, say in the shape of inferential patterns—governing their 'behavior'. Combining these constants with the terms of natural language, we get what we call "hybrid" statements: They are hybrid in the sense that they consist partly of natural language expressions in their raw form and partly of artificial expressions, whose meanings are made explicit in a 'mathematical' way. But though such sentences do not have, strictly speaking, well-defined meanings (they are not expressions of natural language and hence do not have "natural" meanings, but neither do they have stipulated meanings—constructs of formal semantics—for they are not part of an artificial language), we can usually make some sense of them and talk about their truth conditions and even their meanings. The reason is that the hybrid formulas can also be read as shortcuts for certain cumbersome paraphrases of the regimented sentence in the given natural language.[36]

Let us return to the sentence *All animals have heads* which may—employing the means constituted within well known rectification—be regimented as

(1*) $\forall x(\boldsymbol{Animal}(x) \rightarrow \boldsymbol{Has\text{-}head}(x))$.

where $\forall$ and $\rightarrow$ are constants whose meanings are exactly delimited, whereas ***Animal*** and ***Has-head*** are terms about which we merely presuppose that they inherit the semantic properties of the English expressions *animal* and *has head*. We can see (1*) as just a way of putting down the English sentence

(5) *For every individual it is the case that if it is an animal, then it has a head.*[37]

We should be aware of the fact that the regimentation, as well as the paraphrasing, involves a shift in meaning. While the sentence about every individual is clearly false once there is single individual animal which is (be it by some strange coincidence) headless, hardly anybody would consider (1) as false. Moreover, regimenting both (1) and (5) as (1*) involves another meaning shift: While normal speakers would surely not hesitate to infer the fact that something does have a head from (1), it is not correct to infer $\exists x(\boldsymbol{Has\text{-}head}(x))$ from (1*) (at least in so far as we remain within the framework of classical logic).[38]

Footnote 35 continued
streamlining of the grammar. In fact, it is what Quine (1960, Chap. V) calls *regimentation*; but we want to make a terminological distinction between the very introduction of logical vocabulary and the employment of the vocabulary (already introduced) for the purposes of logical analysis of individual sentences. Hence, we reserve the term *regimentation* for the latter and introduce *rectification* for the former.

[36] Authors addressing logical formalization generally note the role of such paraphrases (viz. Blau's 1977, p. 4ff, "explicit rephrasing"; Sainsbury's 1991, 53f, "recovered argument" or Brun's 2003, p. 198ff., "verbalization test"). However, usually they do not fully reflect the problematic status of the hybrid formulas themselves.

[37] Alternatively, we can go the other way around and assimilate the hybrid formulas not to sentences of natural language, but rather to formulas of a fully formalized language—we can look at the natural language expressions contained in them as having formal denotations (in logic, we suppose that both *animal* and *has head* are true of some entities and false of others, hence we can see them as denoting the respective sets of entities of which they are—actually—true).

[38] Baumgartner and Lampert do not seem to mind these kind of semantic differences though they clearly affect truth conditions of the expressions—while normal speakers would surely not consider (1) true in all situations in which there is no animal, (1*) is clearly true in such situations.

We can see that the steps that lead from natural language to hybrid sentences resulting from a regimentation are far from unproblematic. They, however, can yield a significant payoff. (Though we should keep in mind that the payoff is primarily *practical*; in principle, we could do without the hybrid statements and go directly to logical forms.) The paraphrases are, at least in cases that are not too complex, easily understandable—we are usually able to recognize whether the paraphrases say (roughly) the same thing as the sentences they are to regiment (and hence whether our steps towards formalization proceed in the right direction).[39] At the same time they already bring to light the relevant logical structure.[40]

We can naturally consider an inverse step with respect to regimentation, viz. finding out a natural language sentence to which a given regimented sentence corresponds, i.e. 'translating' the sentence formulated in the hybrid language back into natural language. Let us call such a step *verbaliziation*.

4.2 Abstraction

If we are *formalizing* a natural language sentence, then our engagement of the terms of natural language, to be mingled with logical symbols into hybrid sentences, must be seen as merely an intermediary step. (In fact, it is only the process that starts with a natural language sentence and ends with a sentence form that is properly called formalization.[41]) Finally, we must drop the terms borrowed from the given natural language and replace them by utterly meaningless symbols, which we will call *parameters*. Let us call this step away from the hybrid language *abstraction* (as we abstract from meanings of certain expressions). In our case, we obtain the traditional formalization

(1**) $\forall x(F(x) \rightarrow G(x))$.

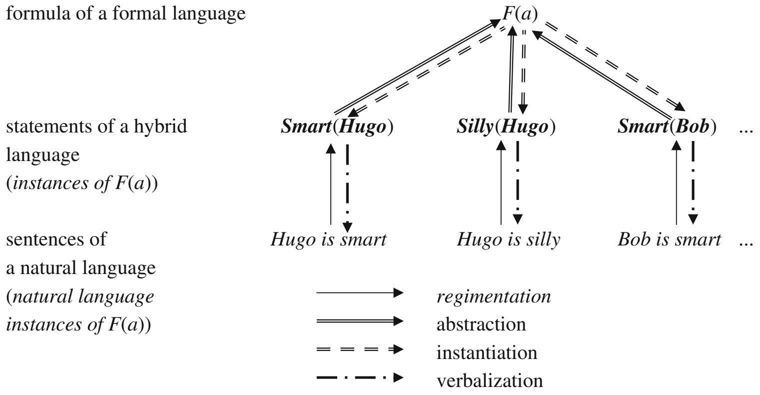
This is no longer a sentence, but a mere *sentence form*, or simply *formula*. The hybrid sentence out of which this formula has been abstracted may be called its *instance*; other instances are all those sentences that result from the replacement of the parameters of the formula by natural language terms of suitable grammatical categories. For simplicity's sake, we will call the natural language sentences that can be regimented by instances of a formula the *natural language instances* of the formula. What is, of course, of the utmost importance is that we take care to represent the same (or synonymous) natural language expressions by the same parameters throughout the whole of the formalized sentence, argument or text.

39 In some cases, we may not rest content with this hybrid language and we may turn it into a fully formalized language, where the semantics of every word is explicitly stipulated and no borrowings from natural language are needed. This is especially possible in case of non-empirical languages, such as the language of Peano arithmetic. (Note that by calling the resulting language *formalized,* we mean that the language is constituted by our definitions; do not confuse this with a language being *formal* in the sense of expressing sentence *forms*. This accords with the terminology of Tarski 1933).

40 In fact, Baumgartner and Lampert's formulas, taken together with the annotations ('interpretations') of their extralogical symbols, can be seen as such hybrid formulas.

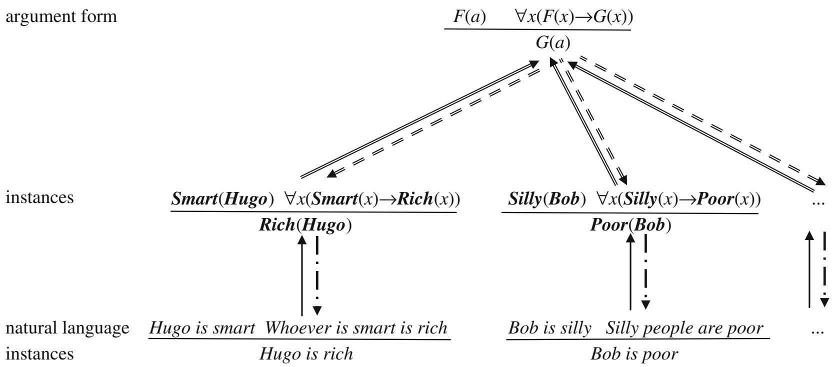
41 As we suggested, Baumgartner and Lampert seem to use the term indiscriminately for regimentation and formalization in this sense.

We can summarize the terminology introduced so far in the following figure:



Formulas can be assembled into *argument forms* in the obvious way. An *instance* of an argument form is what results out of the argument form via a systematic replacement of all parameters by grammatically suitable natural language expressions. A *natural language instance* of an argument form is the result of any verbalization of its instance.

This can be illustrated by the following figure:



If all natural language instances of a formula are (taken to be) *true* (perhaps with some warranted exceptions), we have a good reason to require that the formula comes out as valid in a logical system based on the formal language from which the formula comes (we can also say that the formula is prima facie valid). Similarly, if all natural language instances of an argument form are (taken to be) correct (again perhaps with some warranted exceptions), we can call the form prima facie valid.

5 Criteria of adequacy

5.1 Which formalization we should prefer?

Now we approach the crucial question: If there are two competing regimentations or formalizations of a given sentence of natural language, what are the criteria on the basis of which we decide which one is preferable?

In considering this question we should keep in mind that different logical systems are designed to capture different kinds of inferential relations recognizable in natural languages.[42] These relations are typically associated with the linguistic means that are constitutive of any full-blooded language. Thus, for example, classical sentential logic (CSL) is designed to capture inferential relations associated with the most common sentential connectives; classical predicate logic (CPL) aims, in addition to that, to also capture the inferential relations associated with words like *every* or *some* and with such copulas as *is* or *is not*. The classical modal calculi are designed to also capture (besides capturing the relations captured by CSL or CPL) the inferential relations associated with the classical modal modifiers, etc. Thus, if we wish to use a particular logical system as the means of analysis of a natural language argument, we should be basically clear as to which kinds of natural language inferences should be sanctioned by means of the particular system. If, for example, somebody proposes a logical analysis of the sentences of the argument

John is not a pilot
John is a pilot or Bill is a poet
_____
Bill is a poet

in the language of CSL which classifies the argument as incorrect, we will conclude that his analysis of at least one of the sentences must be incorrect. On the other hand, if he proposes an analysis that will not classify the arguments

John is a pilot
_____
John can aviate

or

John is a pilot
No pilot is an analphabet
_____
John is not an analphabet

as correct, we won't get suspicious in regard to his logical skills concerning analysis in CSL, though intuitively the arguments are correct. To the contrary—if he offers an analysis that will make their correctness directly demonstrable within CSL, we will conclude that he has got his analysis wrong.[43] Of course, often we will not have

[42] We could say that if any system is to be called *logical*, then the inferential relations manifested through the systems should be a matter of *form*; but this could lead us into a vicious circle—the answer to the question which inferential relations are formal is likely to refer to the notion of a logical system.

[43] In fact, in the case of the first of the arguments we would also be suspicious about formalizations in CPL that would treat it as logically correct. Epstein's (2001, p. 167) criterion 2c requiring that "If one proposition

completely clear expectations as to which natural language inferences should be provable as correct within a particular system, but we should have at least a tentative idea as to 'what to expect' from a particular logical system—which natural language arguments can be expected to be *correct* according to the system.

Now let us assume, as a kind of unrealistic thought experiment, that we have at our disposal an adequate formalization of all relatively simple sentences of natural language in the language of a system **S** except for one that will be called B and somebody proposes several formulas $\Phi_1, \Phi_2, \Phi_3, \ldots$ of **S** as candidates for capturing the logical form of B. How should we decide which of them (if any) is the most adequate? The general method is rather simple: We make a list of perspicuous natural language arguments in which B features as a premise or as the conclusion and which we intuitively hold both for correct and falling into the intended scope of **S**.[44]

Alongside this list we make a parallel list consisting of arguments (or 'arguments') composed of the corresponding formulas of **S** in place of the natural language sentences, each entry of the list having alternative variants with different formulas $\Phi_1, \Phi_2, \Phi_3, \ldots$ in place of B. If **S** classifies the formal arguments with Φ_i in place of B as valid only when their natural language counterparts are intuitively correct, then Φ_i passes this test (we can speak of an *adequacy test*) and makes it into a shortlist of candidates for an adequate formalization of B in the language of **S**. (Otherwise, we say that we have discovered a *counterexample*: viz. an incorrect argument which would turn out to be an instance of a valid argument form were we to formalize B as Φ_i.)[45] The longer and more variable the list of testing arguments is, the briefer shall be the shortlist of 'successful' candidates. Finally, we will choose the best candidate(s) from the shortlist on the basis of a loose criteria of simplicity and resemblance of the buildup of the formulas to the buildup (structure) of the sentence B.

If we generalize from the sketch of the method just presented, we can say that the result of the formalization is making explicit the place of B within the inferential structure of its natural language, by means of associating B with a formula of **S** the position of which within the inferential structure of **S** is explicit and definite. Hence **S** serves as a 'map' of the 'inferential surroundings' of B, making it possible for us to gain an overview over its 'inferential landscape' and over the positions of individual sentences and words within it.

Footnote 43 continued
follows informally from another, or a collection of other propositions, then its formalization should be a formal semantic consequence of the formalizations of the other(s) (relative to the logic we adopt)" is, in our view, too general. It is worth noting that, through 2c, inferential considerations sneak into Epstein's approach that otherwise seems clearly semantic.

44 An argument counts as perspicuous if its (in)correctness is intuitively recognized by a normal competent speaker.

45 We, however, should not assume that the appearance of a counterexample unavoidably disqualifies a given formalization. In some cases we do tolerate counterexamples; for example, if we conclude that the tolerance is a price that must be paid for the overall simplicity and perspicuity of the analysis. Thus, for example, we do tolerate the fact that the step from A to $B \rightarrow A$ is an argument valid by classical logic, though many instances of this form (where the implication is instantiated as *if* … *then* …) are intuitively incorrect arguments.

5.2 Criteria

If we try to pin down the inferential (sub)structures of a natural language that we want to make explicit, we will necessarily end up with a slightly fuzzy and gappy network of relations among sets (or sequences) of sentences (premises) and individual sentences (conclusions). The inferential structure of **S** will be, on the other hand, definite, determinate and much simpler. It is not only because the latter is not supposed to be a strictly faithful copy of the former (but rather its idealization, doing away with fuzziness and filling gaps), but also because **S** will be designed to only capture a certain layer of the inferential structure of the underlying language.[46]

In contradistinction to our thought experiment, the formal counterparts of natural language sentences in **S** are hardly ever given so firmly that no revision would be possible and so we, strictly speaking, never really face the task of regimenting a single sentence (the regimentations of all its inferential neighbors being already firmly fixed). Though in the cases of some simple sentences there might not be much leeway, the more complex the grammatical structure of the regimented sentence is, the less clearly determined the regimentation will be. In general, we will have to tolerate the fact that—with perhaps some simple exceptions—any regimentation will always be only tentative even if the particular regimentation successfully passes numerous adequacy tests without any failures.

The problems we face are of two kinds—the first one is that the result of any adequacy test is dependent not only upon the choice of the particular formula that is tested, but also on the formulas selected as formalizations of other sentences appearing in the testing arguments (so we always test a kind of holistic structure while seemingly testing a single formalization). The second problem is that the intuitive classification of the natural language arguments that serve as our reference points is nearly always unavoidably provisional and revisable.

But let us disregard these problems for the time being (we will come back to them later) and instead continue with the assumption that the logical system we are using for our analysis is fixed and that we take the relation of all sentences, save the one, to their natural language counterparts for granted (i.e. that we do not dispute the facts regarding the regimentations of the sentences resp. regimentations of the formulas). If we want to formulate, under these assumptions, the most fundamental criterion of adequacy of logical analysis (formalization) of a sentence, we could perhaps put it in a way that is not very much different from Baumgartner and Lampert's (COR)[47]:

(REL) Φ counts as an adequate formalization of the sentence B in the logical system **S** only if the following holds: If an argument form in which Φ occurs as a premise or as the conclusion is valid in **S**, then nearly all its perspicuous natural language instances in which B appears as a natural language instance of Φ are intuitively correct arguments.

[46] As is the case with any map, the depicting is schematic and simplified, and some aspects of the original are omitted.

[47] Brun (2003, p. 221) discusses an analogous criterion under the name (GK).

We call this *the principle of reliability* as it should guarantee that logical formalization will be reliable in the sense that we can rely on any conclusion drawn by means of logic from accepted premises. It is obvious that (REL) is not a clear-cut criterion. Not only are the concepts of intuitive perspicuity and correctness of natural language arguments loose and open to questioning, but the formulation also contains the inherently problematic phrase "nearly all". However, in our view this vagueness is unavoidable. We simply have to acquiesce to the fact that the parallelism between natural language and its adequate formalization is not to be expected to be perfect—natural language is *inherently* fuzzy and open-ended. In fact, it is the *point* of using the means of formal languages to make things simpler and more orderly.

The obvious complementary (but looser) criterion is the following:

(AMB) Φ is the more adequate formalization of the sentence B in the logical system **S** the more natural language arguments in which B occurs as a premise or as the conclusion that are intuitively perspicuous and correct in the way falling into the intended scope of **S** are instances of valid argument forms of **S** in which Φ appears as the formalization of B.

We call this principle *the principle of ambitiousness*, as it suggests that we cannot make do with an attitude of "the safer the better".[48] Note that this principle is *not* dual to the previous one, viz. it does *not* require that an adequate formalization would have to capture all intuitively valid arguments. Such a requirement would be, as Blau (1977, p. 4) notes "entirely unrealistic".[49] However, it seems that within a definite formal framework, capturing more correct arguments as logically correct is prima facie a virtue of a formalization.

If we wish to provide a truly comprehensive set of criteria, we should add some principles guiding the choice for the cases undecided by the previous criteria. They can be called *the principle of transparency* and *the principle of parsimony*.[50] Their thorough discussion would exceed the limits of this paper, so let us give them only in a very generic form:

(PT) Φ is the more preferable formalization of sentence B in logical system **S** the more the grammatical structure of Φ is similar to that of B.[51]

48 We prefer these names to Baumgartner and Lampert's *correctness* and *completeness*, because the latter might suggest that the relationship between a formal language and a natural language is akin to that between a proof-theoretical specification of theorems and a model-theoretical specification of tautologies, which we find misleading.

49 See also Brun (2003, p. 215) for an argument against this.

50 The first principle is mentioned by Baumgartner and Lampert (2008, p. 112) as part of the traditional picture of adequate formalization (they speak about the formalization being "faithful to the syntactic surface of the respective text" (p. 95)). As a result, their (ADS) corresponds to our (PT). See Brun (2003, §12.1) for a thorough discussion of various specifications of (PT).

51 Assessing similarities of the structures is a delicate matter and it is virtually impossible to formulate anything close to well-determined criteria; but there will be a few clear cases in which the grammatical structures correspond to each other quite straightforwardly, and a not so few cases in which the structures will be alike in important respects.

(PP) Φ is the more preferable formalization of sentence B in logical system **S** the more it is parsimonious as concerns the number of logical symbols it employs.[52]

The import of the principles should be seen as decreasing in the order in which they have been presented. The first of them is sine qua non matter (though even here we can imagine cases in which we allow for exceptions, if they make for a formalization exceptionally suitable from the viewpoint of the other principles). The second is essential as well, as it suggests that the analyst should not search just for 'the safest' formalization but also for the inferentially most 'fruitful' one—the one that makes explicit more relevant valid inferences than competing ones. The last two principles are more-or-less auxiliary (though they can be given more weight within analyses made for certain specific purposes).[53] Thus, especially in the case of the last three, there might be various trade-offs (we might, for example, want to have a regimentation that is not quite transparent if it is exceptionally parsimonious).[54]

5.3 Bootstrapping

Now we must return to the unrealistic assumption we made when we started to look for the criteria of adequacy of formalization, viz. the assumption that we can always see the regimentations of all the sentences except the one whose regimentation we are pondering as fixed. Taken literally it would, of course, lead us into a vicious circle: If we had to base the regimentation of any sentence on already accomplished formalizations of other sentences, the whole enterprise would never really be able to get off the ground.[55]

The solution, of course, is bootstrapping: We start with mere tentative regimentations of some simple sentences, basing the regimentations of others on them. Hence, if we are considering Φ as a possible regimentation of B and we find out that some argument form involving Φ as a counterpart of B is valid, whereas there is a counterexample, we will not only consider dropping the hypothesis that Φ is an adequate formalization of B, but will also take into account the possibility of keeping the

[52] We are aware of the fact that this formulation is ambiguous since we can count the number of *types* of logical symbols in a given formula, as well as the number of *occurrences* of logical symbols in the formula. The ambiguity is benign, as the principle in fact requires parsimony of both kinds.

[53] Blau (1977, pp. 18–19) proposes to complement the semantic criterion by a single additional one stating that the formalizations are "produced by the simplest and most general transformational rules possible" (cf. also Brun 2003, p. 269). It is, however, unclear whether there is anything such as *the* simplest and most general rules; and whether excessive striving for simplicity and generality could not lead us to an intolerable deviation from the structure of natural language. The principle (PT) is to block this, whereas the principle (PP) compensates for the resulting pull to the surface structure of natural language by the demand of minimization of the resources on which formalization draws.

[54] In certain contexts, (PT) might be imperative: if we are, for example, building an AI system that is to have natural language specifications of problems as its inputs, then we may want to work with logical regimentations as close to natural language as possible (to make the process of regimentation easy for the system). If, on the other hand, what we are after is searching for "the smallest store of materials with which a given logical or semantic edifice can be constructed" (Russell 1914, p. 51), then it will be probably rather (PP) that we will find imperative.

[55] Baumgartner and Lampert call this the *circularity problem* (p. 97).

hypothesis at the cost of dispensing with formalizations of some of the other sentences involved in the counterexample. Thus, the formalization is in fact a holistic, give-and-take enterprise.

Another problematic assumption that we should reconsider at this point is the assumption that we start the regimentation with a fixed formal language and hence with a fixed stock of logical constants. By adopting this assumption, we adopt an 'internal' perspective on the process of regimentation (and formalization): a perspective from inside of a given logical system. This might have been a perspective that appeared as general to some pioneers of logical analysis—who took the logic they had as the only one possible, or at least the only *true* logic. We are now, however, well aware of a wealth of alternatives. In addition, and this is crucial for us now, it is not clear as to where one particular logic could acquire the normative force which would allow it to discard all its alternatives.[56]

We think that the only possibility of vindicating something as logic in the narrow sense of the word, and in any case as a reasonable tool of logical analysis, is its ability to reflect the 'logical structures' of natural language, especially envisaging its inferential interlinking. Hence, any logical vocabulary we might use as a tool of analysis is dependent, for its status as a *logical* vocabulary and as a *proper* tool of logical analysis, on its being useful in forming a map of natural language.

Let us consider a concrete symbol of common logical vocabulary—implication ($\to$). One of the rules which characterizes its behavior within common logical systems is *modus ponens*

$$\frac{A \quad A \to B}{B}$$

Given that $\to$ is the result of the well-known rectification of the English *if-then* phrase, this rule manifests the fact that most of the English arguments of the form

$$\frac{A \quad \mathit{If}\, A, \mathit{then}\, B}{B}$$

are generally accepted as correct. Not *all* of them must be such (and probably not all of them are); it is enough that the number of those accepted as correct warrants taking the rest as deviations, which are *not really* arguments of this form.

Consider the following example:

$$\frac{\mathit{A\ farmer\ owns\ a\ donkey} \quad \mathit{If\ a\ farmer\ owns\ a\ donkey,\ then\ he\ beats\ it}}{\mathit{He\ beats\ it}}$$

Despite the fact that it can be seen as a natural language instance of the above form, it is at least dubious that it is a sound argument—the trouble is that its conclusion (unlike *Pedro beats Chiquita* or *A farmer beats a donkey* or *Every farmer who owns a donkey*

[56] Once you put up with the fact that there is a spectrum of logical systems to choose from, you may assume an 'external' perspective, weighing pros and cons of individual systems available as tools of logical analysis.

beats it), is somewhat incomplete.[57] Hence the defender of *modus ponens* will have to come up with an explanation, which, in this case, is not difficult to make up: We would say that A and B cannot be replaced by any English sentence whatsoever, but rather only by sentences which are in some sense 'self-contained', which 'express a proposition'.

In this sense, the rule is normative (and logic becomes not merely *ars explicandi*, but also partly an *ars iudicandi*): As long as logical rules are in force, they 'force us' to classify some initially plausibly looking arguments as unacceptable or to explain away some arguments which look like instances of certain schemata as 'in fact' their non-instance. But once such explanations start to be too cumbersome we might decide to give up on a rule. Hence, here we have one more, and a more basic give-and-take. This is the reflective equilibrium at work.

6 Summing up

It can be seen that our approach differs from that of Baumgartner and Lampert in several respects.

(1) We reject an atomistic conception of regimentation and argue that formalization is inherently a holistic endeavor. (The only way to put constraints on formalization of a particular sentence is to fix formalization of other sentences, so the formalizations inevitably create mutually supporting network.)

(2) We reject the view that the semantic criterion of an adequate formalization based on truth conditions is preferable to the inferential. If we proceed through regimentation of the natural language sentence then, naturally, semantic considerations can be useful. But, in general, every aspect of the semantic criterion that can be accounted for at all is already contained in the inferential criteria; relying on truth conditions may be misleading.

(3) We reject the contextualist conception of regimentation according to which the same sentence is (typically) to be regimented differently within different texts (though we believe that we can legitimately reach different regimentations when doing logical analysis pursuing different goals).

(4) We reject the view that an adequate *logical* formalization should capture all inferential links—the logical ones as well as the material ones. On the level of particular sentences and arguments we should always see formalization as an endeavor based on a particular logical language (or system) designed to capture only inferential relations of certain kinds (the ones that the particular language is designed to capture).

(5) We reject the view that formalization should aim at a *perfect* map of the 'logical landscape' of natural language. Even formalizations that are far from perfect may be adequate for certain purposes; indeed, some amount of streamlining and idealization is part and parcel of making maps.

[57] We feel that the conclusion should be detachable from the argument, but saying that we have proven that "he beats it" does not seem to make much sense.

(6) We do not think that criteria that can be revised on the basis of considerations of the reflective equilibrium are no criteria at all (but mere "rules of thumb", as Baumgartner and Lampert put it)—we are convinced that *any* humanly usable criteria are in principle revisable (though, of course, the revision may only result from some essential and global discrepancies, not from some marginal and local ones).

(7) Though we agree with Baumgartner and Lampert that formalization can be construed as *explication*, we disagree that it is not an explication in the sense of Carnap. We propose to understand it precisely in the Carnapian sense[58]: As a replacement of something essentially vague and open-ended with something exactly and explicitly defined, doing away with the vagueness and closing up the open ends. (It can be also seen as what Wittgenstein called *übersichtliche Darstellung*, i.e. *perspicuous representation*.[59])

This list might seem to indicate that there is almost nothing in Baumgartner and Lampert's paper with which we would agree. But such an impression would be mistaken. In fact, we think we agree with them on many crucial issues regarding the very nature of logical analysis. We agree that assessing the success of a logical analysis is not a matter of considerations concerning some metaphysical substratum of what we say, but rather of down-to-earth criteria of assessment of the proximity of the result of the regimentation resp. formalization with the part of natural language that was its point of departure. We agree that in order to look for the criteria we must start from the considerations concerning the match between the explicit inferential structure of the regimenting language and the pretheoretically given structure of the regimented one, and that if this is not enough, we should consider the closeness of surface forms (though we reject the assertion that we must ultimately "go semantic"). We agree that logic may be seen primarily as *ars explicandi* (in Baumgartner and Lampert's sense, i.e. in the sense that it accounts for the immanent patterns of our argumentation and reasoning, rather than spelling out some transcendent rules of human reason), but we think that the explication is inherently connected with introducing standards, and in this sense logic inevitably acquires the status of *ars iudicandi*.

We are also prepared to admit that our conception of adequate formalization does not achieve the goals that Baumgartner and Lampert wanted to achieve (and claim to have achieved). The reason is that we consider their goals to be unrealistic. The principles of an adequate formalization that we present are not intended to lead us to *the* (only) adequate formalization, they are just meant to give guidance for comparing different formalizations and deciding (perhaps not always unequivocally) which of them are preferable (with respect to a given purpose). So we don't take the problems invoked by Baumgartner and Lampert—viz. the termination problem or the circularity problem—too seriously. As we suggested, logical formalization is an activity

[58] Carnap (1950, p. 3) characterized his notion in the following way: "By the procedure of explication we mean the transformation of an inexact, prescientific concept, the explicandum, into a new exact concept, the explicatum. Although the explicandum cannot be given in exact terms, it should be made as clear as possible by informal explanations and examples. The task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second."

[59] See Wittgenstein (1953, §122). Cf. also Peregrin (1992, 2001).

with outcomes of two kinds. Firstly, it improves our insight into (and control over) our inferential practices; secondly, it serves to substantiate the fact that some formal systems are logical and as such deserve special attention (and scrutiny in journals devoted to logic).

7 Conclusion

The crucial upshot of our considerations should not be seen in the comparison of our approach with that of Baumgartner and Lampert. Rather, it should be seen in the defense of the fundamental role of reflective equilibrium considerations in establishing any logic worth the name, and in understanding the role of logic in relation to natural language.

The laws articulated by logic are not a mere reflection of something that exists, in a wholly articulated shape, either within our thinking or somewhere under the surface of our language. There is no way of merely extracting already completed laws of logic directly from there—what we can get as the starting point of logic are certain patterns of valid inferences that are accepted across different domains of our discourse and reasoning, but which are not quite definite (both in the sense of not being exceptionless and in the sense of not having an utterly clear-cut semantics).

This implies that any kind of logical system may only partially be based on patterns which logicians simply *find* and *report*—it must be *also* be based on *completions* and *streamlinings* that logicians perform. Hence, the laws of logic as they have been articulated by logicians, though crucially reflecting pre-existing patterns of valid inference, go well beyond them. Thanks to this and thanks to the—modest but extant—attention that the work of logicians receives, logic influences the language of science and consequently even—slightly—the colloquial idiom, and comes to be taken as a *norm*. It acts as a norm of what is to be seen as regular and what as 'irregular', what is a lawful usage and what an exception. (In this way it secures a framework for adjudicating various disputes that would otherwise be hardly resolvable.)

We have tried to portray how this works in terms of the dialectics of correct inferences and valid forms. Some inferences (in natural language) are prima facie correct, which makes us see some forms of inferences (namely those which have correct instances) as prima facie valid. However, we take the quest for (getting a grasp on) validity as an instance of a quest for *e pluribus unum*, as a quest for finding a perspicuous order within the unorderly vastness of individual correct and incorrect inferences; and this makes us impose more order on our language and our reasoning than we are able to *find* there, even at the cost of some Procrustean trimming and stretching. Hence, upon reflection, a form of inference comes to be taken as valid, not literally in cases when all their inferences are correct, but in cases when those which are not can be reasonably explained away.[60]

More traditional approaches to logical formalization often create the illusion that behind or beneath the surface form of our language there is some deeper and more

[60] In this way, some prima facie correct inferences may be demonstrated as incorrect using the notion of validity which was originally based on this correctness.

substantial logical form. However, we do not believe that we can get to such a form by a process substantially different than the 'give-and-take' one described above, hence by a process led by the maxim of simplicity and maximal order—the maxim that is operative in any science. In particular, we do not believe that we can get from the surface form to the logical form by some process that has nothing to do with reflective considerations and hence logic, so that logic would then be left with the task of pulling out the ready-made structure and lending it a perceptible form. We are convinced that the way from the surface to the so-called logical form involves considerations largely constitutive of logic, so that the resulting logical form is not what logic merely describes or reports, but rather what logic helps bring into being.

All of this implies that though (a) logic cannot derive its authority (and usefulness) from anything else than the existing patterns of our reasoning (especially inferential patterns of our language that are taken to be correct), (b) such patterns do not determine logic uniquely, hence, (c) logic draws also on maxims of simplicity and order, and (d) this surplus contribution allows for adopting the view that logic assumes a normative position over our language and reasoning and tells us what is correct and what not. This, we think, is the essence of the reflective equilibrium (in the present context) and it is, we are convinced, the correct way to understand the position of logic with respect to our language and our reasoning.

Acknowledgments Work on this paper was supported by research grant No. P401/10/1279 of the Czech Science Foundation. We are grateful to Georg Brun and the referees of the journal for their detailed and valuable critical comments.

References

Baumgartner, M., & Lampert, T. (2008). Adequate formalization. *Synthèse, 164*, 93–115.
Blau, U. (1977). *Die dreiwertige Logic der Sprache*. Berlin: de Gruyter.
Blau, U. (2008). *Die Logik der Unbestimmtheiten und Paradoxien*. Heidelberg: Synchron.
Brun, G. (2003). *Die richtige Formel*. Frankfurt: Ontos.
Carnap, R. (1950). *Logical foundations of probability*. Chicago: University of Chicago Press.
Epstein, R. L. (2001). *Propositional logics* (2nd ed.). Belmont: Wadsworth.
Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.
Peregrin, J. (1992). Sprache und ihre Formalisierung. *Deutsche Zeitschrift für Philosophie, 40*, 237–244.
Peregrin, J. (2001). *Meaning and structure*. Aldershot: Ashgate.
Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
Resnik, M. (1985). Logic: Normative or descriptive? The ethics of belief or a branch of psychology?. *Philosophy of Science, 52*, 221–238.
Russell, B. (1914). *Our knowledge of the external world*. London: Allen and Unwin.
Sainsbury, R. M. (1991). *Logical forms (an introduction to philosophical logic)*. Oxford: Blackwell.
Svoboda, V., & Peregrin, J. (2009). *Od jazyka k logice [From language to logic]*. Prague: Academia.
Tarski, A. (1933). Pojęcie prawdy v językach nauk dedukcyjnych. Warsawa (English translation: The concept of truth in formalized languages in Tarski, A. (1956). *Logic, semantics, metamathematics* (pp. 152–278). Oxford: Clarendon Press).
Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Routledge (English translation: London: Routledge, 1961).
Wittgenstein, L. (1953). *Philosophische Untersuchungen*. Oxford: Blackwell (English translation: *Philosophical investigation*, Oxford: Blackwell, 1953).