Philosophical Studies (2024) 181:1067–1085
https://doi.org/10.1007/s11098-024-02120-y

Logicality in natural language

Gil Sagi[1]

Accepted: 2 February 2024 / Published online: 29 April 2024

Abstract
Is there a relation of logical consequence in natural language? Logicality, in the philosophical literature, has been conceived of as a restrictive phenomenon that is at odds with the unbridled richness and complexity of natural language. This article claims that there is a relation of logical consequence in natural language, and moreover, that it is the subject matter of the bulk of current theories of formal semantics. I employ the framework of *semantic constraints* (Sagi in Log Anal 57(227):259–276, 2014), which generalizes the Tarskian definition of logical consequence. I apply the widely accepted criterion of invariance under isomorphisms (Sher in J. Symb Log 61(2):653–686, 1996) generalized to the framework of semantic constraints (Sagi in Bull Symb Log 28(1):104–132, 2022b), combined with a theory of Glanzberg (in Metasemantics: new essays on the foundations of meaning, 2014) to delineate the relation of logical consequence in natural language.

Keywords Logical consequence · Logicality · Natural language semantics · Invariance criteria

1 Introduction

Is there a relation of logical consequence in natural language? How should we understand this question? How can we make it interesting? I believe that there are several different interesting ways of understanding the question, involving different interpretations of "logical consequence", "natural language" and logical consequence being "in" natural language. Here I would like to explore one of them.

In a recent article, Glanzberg (2015) gave a negative answer to the question of logic in natural language. Here, my understanding of the question will be very much

✉ Gil Sagi
gsagi@univ.haifa.ac.il

[1] Philosophy Department, University of Haifa, 199 Aba Khoushy Ave., 3498838 Mt. Carmal, Haifa, Israel

in line with Glanzberg's, but with some divergences that will lead to an opposite conclusion.[1]

Natural language will be treated here as a natural phenomenon—as the object of study of empirical linguistics. Logical consequence will be taken to be a relation between sets of sentences (constituting premises) and sentences (serving as conclusions) in a given language. This relation holds if the conclusion necessarily follows from the premises by virtue of the form of the sentences. These assumptions are widely shared, though developing the understanding of form beyond common contentions will be one of the present objectives. Both logic and natural language can be seen to involve both normative and descriptive aspects. Here, we'll put normative issues aside: the approach to natural language will be descriptive, and the formal framework will be a tool used in its study.

As Glanzberg points out, one can be restrictive or permissive to varying extents in one's approach to logical consequence. On a permissive approach, one can surely identify a relation of logical consequence in natural language. Glanzberg's negative conclusion regards the restrictive take on logic. Glanzberg's restrictive take basically limits logic to classical first or second order predicate logic. My approach to logic here will be more permissive than that. Nonetheless, my approach will be in keeping with a common way of restricting logic, countenanced by Glanzberg, using invariance criteria of logicality.

In the present context, when we ask whether there is a logical consequence relation in natural language, one way to approach the issue would be to see whether formal systems that satisfy basic conditions we would expect from systems for logic are good models for some phenomenon in natural language. I shall claim that indeed, contemporary semantic theory for natural language can be phrased as such a formal system. The overarching formal framework that I will use will be a model-theoretic framework of semantic constraints [as in Sagi (2014)], and the criterion by which I shall explicate logicality will be invariance under isomorphisms [as in Sher (1996)].

I will use contemporary theories of formal semantics as our access to linguistic phenomena. The understanding of logicality, on the other hand, will be taken from the philosophy of logic.

Now, if we look at basic examples from semantic theory, we see that despite the extant use of formal tools, the results are far from what would be considered a logic by those interested in drawing bounds for logic. Glanzberg bases his arguments on such examples:

> What is characteristic of most work in the model-theoretic tradition is the assignment of semantic values to all constituents of a sentence, usually by relying on an apparatus of types (cf. (Chierchia & McConnell-Ginet, 1990; Heim & Kratzer, 1998)). Thus, we find in model-theoretic semantics clauses like:[2]

[1] For a critique of Glanzberg (2015), see Sagi (2022a), where I contest Glanzberg's arguments, and argue that his own assumptions even lead to a positive approach towards logic in natural language.

[2] Glanzberg explains:

In common notation, $[\![\alpha]\!]$ is the semantic value of α. I write $\lambda x \in D_e.\phi(x)$ for the function from the domain D_e of individuals to the domain of values of sentences (usually truth values) (*ibid*).

(3) a. $[\![\text{Ann}]\!] = \text{Ann}$

b. $[\![\text{smokes}]\!] = \lambda x \in D_e.\ x \text{ smokes}$

> ... [These clauses] provide absolute statements of facts about truth and reference... We see that the value of ‘Ann’ is Ann, not relative to any model. (Glanzberg 2015, p. 89)

The expressions ‘Ann’ and ‘smokes’ are, by accepted restrictive standards, nonlogical terms. Accordingly, they shouldn’t be fixed in a definition of logical consequence. There might be an intended model for the full vocabulary of the language, but logical consequence is defined on the backdrop of a range of models, abstracting away from the meaning of nonlogical terms. But in natural language, as noted by Glanzberg, we find entailments that defy accepted delineations of the logical-nonlogical vocabulary:

(6) a. We loaded the truck with hay.
ENTAILS
b. We loaded hay on the truck.

DOES NOT ENTAIL

We loaded the truck with hay.

> (Glanzberg 2015, p. 93)

Here, the claim is that these entailments require fixing ‘load’, ‘truck’ and ‘hay’ as logical terms. There is a consensus among all those who consider logical consequence to be a formal relation that these words are not logical. For one, they are not invariant under isomorphisms, taken by many to be at least a necessary condition on logical terms. Previously, I claimed that these examples do not show that there isn’t a relation of logical consequence in natural language—they still leave room for a restricted relation that is a proper subset of all entailments (Sagi, 2022a). Here, I take a different approach. First, I deny that those entailments require fixing ‘load’, ‘truck’ and ‘hay’ as logical. Moreover, these entailments can be captured in a logic that abides by the criterion of logicality of invariance under isomorphisms. In addition, while I agree that the above clauses (3a-b) are nonlogical as they stand (meaning that they defy the isomorphism invariance criterion for logicality), they should be cleaned up in order that we can discern their actual contribution to semantic theory.

The plan of the paper is as follows. In the next section (Sect. 2), I present the framework of *semantic constraints*, which generalizes the standard first order model-theoretic semantics. I explicate the notion of logical consequence through this framework (Sect. 2.1), and generalize the commonly accepted criterion for logicality of isomorphism invariance to apply in the framework (Sect. 2.2). In Sect. 3, I discuss another paper by Glanzberg, “Explanation and Partiality in Semantic Theory” (2014). I show how Glanzberg’s ideas presented there, in combination with the framework of semantic constraints, can give a theory of logical consequence in natural language.

2 The framework of semantic constraints

The logical framework of semantic constraints is a model-theoretic framework that is not based on a sharp division of the vocabulary into logical and nonlogical. In this section I present the framework. In Sect. 2.1, I present the basics of the framework, following (Sagi, 2014). In Sect. 2.2, I show how a criterion of invariance under isomorphisms can be generalized and reformulated for this framework, based on (Sagi, 2022b). Then, in the next section (Sect. 3), I apply the framework and its notion of logicality to semantics of natural language, offering a theory of logical consequence in natural language.

2.1 The basic framework

In his *argument from lexical entailments*, Glanzberg claims that if we were to capture all entailments such as those mentioned on p. 3, we would have to fix 'load' and 'cut' as logical constants, and that if we were to capture all entailments in natural language, we would probably need to fix all expressions as logical (Glanzberg, 2015, §3.3.2).[3] The framework of semantic constraints avoids this conclusion. In this framework, one can fix the meanings of terms in the language in various ways and to various degrees, sufficient to capture some entailments without committing to a determinate reference. We thus lose the strict division of the vocabulary into logical and nonlogical, the former completely fixed and the latter maximally variable. This is a formal framework, but it can be used to model natural language semantics as other model-theoretic frameworks do.

A language L will consist of *terms* (the primitive expressions of L) and *phrases* (the meaningful expressions of L, which include the terms and strings of terms with auxiliary symbols). A *semantic constraint* is a statement in the metalanguage which restricts the range of models for the language. Given a set of semantic constraints, we obtain a class of models, and we define a logical consequence relation as truth preservation over that range of models. Clauses fixing logical terms are a special case of semantic constraints: they restrict the range of models to those that give the fixed meaning to those terms.

We should say what models are here, as they are more general than in the standard model-theoretic framework. We start with a very wide class, which is then narrowed down using constraints. A *model* is a pair $\langle D, I\rangle$ where D is a non-empty set (the *domain*), and I an interpretation function which assigns to phrases in L values in the set-theoretic hierarchy built over $D \cup \{T, F\}$ (T and F are the truth values). The function I can assign any such value to any phrase of L, and it becomes limited only by adopting a set of semantic constraints.

The usual semantic clauses for connectives and quantifiers can be formulated as semantic constraints—those are special cases which completely fix the extension of

[3] Glanzberg qualifies this strong conclusion by considering the option that certain patterns among families of expressions can be identified that will allow us not to take every single word as a constant (Glanzberg 2015, p. 95), but still, nothing like a logic (according to Glanzberg) would come out.

the terms. For example, '$I(\exists) = \{A : \emptyset \neq A \subseteq D\}$' is a semantic constraint which limits us to the class of models where the existential quantifier receives its standard interpretation as a second level predicate which holds of all and only non-empty sets. Then, we have a semantic constraint such as '$I(Red) \cap I(Green) = \emptyset$' which fixes neither *Red* or *Green*, but constrains their interpretations to be mutually exclusive. Then there will be semantic constraints regarding the interpretation of non-atomic phrases (that may cover, for example, standard recursive clauses). For examples and further details, I refer the reader to Sagi (2014, 2022b).

As formulated here, the framework is extensional, and the models are just a domain and an interpretation function. But surely, the framework can be extended to an intensional setting, where models include a range of possible worlds. Current semantic theory, where we would like to apply the framework of semantic constraints, is normally phrased using intensional semantics. However, since the examples we shall deal with do not involve intensional operators, we can continue to use extensional models for the sake of simplicity. And so, we can reformulate the semantic clauses presented in the previous section as semantic constraints, adding a third that will serve us later on.

1. (a) $[\![\text{Ann}]\!] = \text{Ann}$
   (b) $[\![\text{smokes}]\!] = \lambda x \in D_e.\, x \text{ smokes}$
   (c) $[\![\text{most}]\!]_M = \{\langle A, B\rangle \in \mathcal{P}(M)^2 : |A \cap B| > |A \backslash B|\}$

   will be reformulated as:[4]

2. (a) $I(\text{Ann}) = \text{Ann}$
   (b) $I(\text{smokes}) = \lambda x \in D.\, x \text{ smokes}$
   (c) $I(\text{most}) = \{\langle A, B\rangle \in \mathcal{P}(D)^2 : |A \cap B| > |A \backslash B|\}$

The question now is whether there is a way to demarcate the "logical" semantic constraints, those that would provide the logic of natural language. We give an answer in the next subsection. Before that, we mention some reasons from linguistics for moving to a generalized conception of logicality provided by Chierchia (2021). The distinction between logical/nonlogical terms, or rather between function/content items has undeniable theoretical significance in linguistics (although, how exactly these distinctions line up is itself a theoretical issue). Logicality, as argued by Chierchia and others, is intricately connected with grammaticality, and more specifically, logical triviality explains cases of ungrammaticality (Gajewski, 2002; Chierchia, 2013, 2021; Del Pinal, 2019). Criteria for logicality are thus relevant also for sheer linguistic concerns. And those concerns too may invite the kind of view we present

[4] Note that constraint 2(a) fixes the extension of 'Ann' to be Ann, and by that restricts us to models where Ann is a member of the domain. We read constraint 2(b) as giving to 'smokes' as semantic value the function, from the domain to the truth values that for every element of the domain gives the value T if and only if that element smokes (see n. 2).

here. Chierchia gives the example of gender features, with the following semantic clauses from Italian:

(37) a. i. $||\text{fem}|| = \lambda x_e$: female($x_e$). x_e $\quad ||\text{male}|| = \lambda x_e$: male($x_e$). x_e
ii. $||\text{ragazz-a}|| = \lambda x_e$: fem($x_e$). young adult($x_e$)
iii. $||\text{ragazz-o}|| = \lambda x_e$: male($x_e$). young adult($x_e$)

b. $\forall$x [female(x) → ¬male(x)]

...

> Use of features of this sort induces disjointness constraints such as (37b), which are among the most common across languages. This seems to require an extension of what counts as 'logical' to constraints that define 'subcategories' of various content words. (Chierchia 2021, p. 247)

The simplicity of standard "term-based" logical systems is certainly a virtue, yet such systems are just a special case from the perspective of the framework of semantic constraints—logical consequence does not *eo ipso* require a strict division of the vocabulary into logical and nonlogical. This applies also, more specifically, to natural language semantics: term-based systems have been useful in its study, but we shouldn't be limited to them when considering logical consequence in natural language. Chierchia's example motivates forgoing the familiar term-based perspective.

2.2 Invariance criteria in the framework

The framework of semantic constraints is a generalization of standard model-theoretic semantics: in it, first order predicate logic is one system among many, where semantic constraints can be added to it or removed from it. Formality here receives a new interpretation. On the standard conception, the form of an argument is determined by the logical vocabulary. Logical terms provide the skeletons of argument schemas, and an argument is valid if and only if it is an instance of a schema all of whose instances are valid. Here, forms of arguments are rather determined by semantic constraints. Accordingly, the notion of a *schema* receives a new interpretation (for details, see Sagi, 2014, p. 270).

When it comes to the question of the logicality of natural language, the demarcation of the logical vocabulary is regarded as a central issue. Standard logical systems include a strict demarcation of the logical vocabulary, and that is the vocabulary that has semantic clauses fixing its interpretation. One of the guiding questions in constructing such systems is which terms are logical, and so will have their interpretation fixed. We have mentioned invariance under isomorphisms as a widely accepted criterion for logicality.

In the framework of semantic constraints, the question of the logical vocabulary loses its significance: now the choice is not of a subset of the vocabulary that gets a fixed interpretation, but rather more generally of constraints on models, restricting the interpretations of terms but not necessarily fixing them completely. In Sagi (2022b) I present a generalization of invariance criteria for logical terms that applies to semantic constraints, including those that completely fix the interpretation of terms and those that do not. Here, I shall merely phrase the generalization, and see how it may apply to natural language semantics, and thus to the question of logical consequence in natural language.

The idea of the generalization is the following. A semantic constraint will satisfy the condition of invariance under isomorphisms if it does not distinguish between isomorphic models. I shall formulate the invariance criterion for both terms and constraints below. But first let me mention the primary motivation for the invariance criterion for logical terms: logical terms should not distinguish between elements in the model domains; logical terms should be blind to permuting elements or switching them with others. The same idea will apply to semantic constraints: a semantic constraint satisfying the condition will be indifferent to permuting or switching elements of domains with others.

Both criteria will refer to bijections between model domains, extended to truth values and sets formed over the domains. We thus define recursively, for any function f from a set D to a set D' the function f^+ on elements in the set-theoretic hierarchy over $D \cup \{T, F\}$ as follows (we assume that $T, F \notin D \cup D'$ and also that D and D' consist of *ur-elements*, and so they include as members no sets built over D and D', so that a recursive definition can be applied):

- $f^+(x) = \begin{cases} f(x) & \text{if } x \in D \\ x & \text{if } x \in \{T, F\} \end{cases}$

and for a set A belonging to the set-theoretic hierarchy over $D \cup \{T, F\}$,

- $f^+(A) = \{f^+(B) : B \in A\}$

Since this extension of f is a natural one, we will omit the superscript and simply speak freely of f applying to the relevant sets.

Now, when a term t is claimed to be invariant under isomorphisms, what is standardly meant is that there is an operation O_t associated with t that gives its intended interpretation in all domains, and this operation is invariant under isomorphisms. And so, if we accept invariance as a criterion for logicality, we may then fix t in all models by the operation O_t. We note that the standard term-based framework assumes that each candidate logical term has some such operation associated with it (either stipulated or capturing a preconceived meaning), and this assumption, required by invariance criteria for logical terms, is not needed in the framework of semantic constraints.

Definition 1 (*Invariance under isomorphisms: terms*) Let t be a term and O_t be the operation associated with t. The term t is *invariant under isomorphisms* if for any sets D and D' and a bijection $f : D \to D'$ appropriately extended, $f(O_t(D)) = O_t(D')$.

The standard logical terms of first-order logic are invariant under isomorphisms. In addition, so are generalized quantifiers such as *Most* and $\exists_{\aleph_0}$ (*there exist infinitely many*).

The definition of invariance under isomorphisms for semantic constraints uses the notion of isomorphic models, which we define below. Note that since, in this framework, we do not assume a division of the vocabulary into logical and nonlogical, or recursive clauses for complex formulas at the outset, we cannot use the standard definition of isomorphic models which refers to the interpretations of the nonlogical terms. The definition of isomorphism thus refers to all phrases in the given language.[5]

Definition 2 (*Isomorphic models*) We say that $M = \langle D, I \rangle$ is *isomorphic to* $M' = \langle D', I' \rangle$ ($M \cong M'$) if there is a bijection $f : D \rightarrow D'$ that when appropriately extended yields $f(I(p)) = I'(p)$ for every phrase in the given language L.

And so, isomorphic models have the same structure imposed on them by the interpretation function. The idea of the following definition is that semantic constraints that are invariant under isomorphisms will not distinguish between isomorphic models.

Definition 3 (*Invariance under isomorphisms: semantic constraints*) A semantic constraint C is invariant under isomorphisms if for any models M and M' such that $M \cong M'$, if M is a $\{C\}$-model, then M' is a $\{C\}$-model.[6]

Note that the models M and M' in the definition are *any* models, not necessarily satisfying some given set of constraints. Basically, a constraint on models is invariant under isomorphisms if the class of models satisfying the constraint is closed under isomorphisms. In Sagi (2022b) it is argued that this the right way to generalize the criterion of invariance under isomorphisms for logical terms. In the special case of semantic constraints fixing terms completely, the criteria are equivalent—it's proved that a term is invariant under isomorphisms if and only if its associated semantic constraint fixing its interpretation is invariant under isomorphisms. Moreover, the generalized criterion is faithful to the desideratum of not distinguishing between individuals. We can thus formulate the following criterion (to be modified later on):

> *Invariance criterion for semantic constraints*: A semantic constraint is logical if it is invariant under isomorphisms.

Now, which semantic constraints are invariant under isomorphisms? A whole host of them. Any semantic constraint that does not distinguish between individuals in the domain vis-á-vis their non-semantic properties will be invariant under isomorphisms. Some examples include:

[5] For the relation between the present definition of isomorphic models and the standard one, see Sagi (2022a, p. 127).

[6] Cf. Zimmermann (2011, p. 790).

- $I(\exists) = \{A : \emptyset \neq A \subseteq D\}$[7]
- $I(Red) \cap I(Green) = \emptyset$
- $I(Red) \subseteq I(Colored)$
- $I(wasBought) = I(wasSold)$

And so, if we use the criterion of invariance under isomorphisms to explicate logicality, but generalize to a framework of semantic constraints, logic appears to be much more permissive than in the special case of a term-based semantics. However, not all semantic constraints pass the test. Examples of semantic constraints that are not invariant under isomorphisms include:

- $I(naturalNumber) = \{0, 1, 2...\}$
- $3 \in I(Odd)$
- $I(Even) \cap I(Prime) = \{2\}$

From these examples we see that semantic constraints that pertain to the "material" of the domain, to what the domain is made up of, are those that are not invariant under isomorphisms.

Still, the few examples we give seem to support the impression that the criterion of invariance under isomorphisms vastly overgenerates. Indeed, any constraint stating a cardinality property or a set-theoretic relation between the interpretation of terms will be invariant under isomorphisms, and so the following semantic constraints satisfy the criterion as well:

- $I(John) \in I(Bachelor)$
- $I(Red) \cap I(Big) = \emptyset$
- $|I(Red)| \geq 375$ (i.e., the size of the extension of *Red* is at least 375.)

These are semantic constraints we would not like to accept in a theory for natural language—of either its logical or its nonlogical part.[8] Surely, invariance under isomorphisms is not a sufficient condition for accepting a semantic constraint. The problem here seems to be that the example constraints are not faithful to the intended meanings of the terms involved, and in any case do not set conditions that would be reasonable to accept in semantic theory.[9]

The line I will take here will be based on a proposed strengthening of the invariance criterion for logical terms by McGee. McGee (1996) supports permutation invariance as a criterion for logical *operations*, and submits that a criterion for

[7] Recall that semantic constraints implicitly generalize over domains and interpretation functions: 'I' stands for an interpretation function, 'D' stands for a domain.

[8] It may seem odd to consider these constraints as "semantic", but all that is meant by the term is that these are restrictions on the range of models.

[9] In the case of the invariance criterion for logical terms, examples along similar lines have been presented in Gómez-Torrente (2002). For even further examples regarding invariance and meaning, see McCarthy (1981), Hanson (1997), for a response, see Sagi (2015) and for further proposals and discussions, see Woods (2017), Zinke (2018).

logical *terms* cannot be straightforwardly derived: while the denotation of a logical term must be permutation-invariant, permutation invariance does not suffice, and meaning in a wider sense than mere extension comes into play. McGee gives a starting point for the discussion on the logicality of terms by posing the following as a conjecture:

> A connective is a logical connective if and only if it follows from the meaning of the connective that it is invariant under arbitrary bijections. (McGee, 1996, p. 578)

McGee admits that what it means for invariance to follow from the meaning of the connective is somewhat vague. We can offer the following explication. In the present context, we take semantic theory for natural language to decide what follows from the meanings of expressions. Further, we are interested in a criterion for semantic constraints rather than connectives. So our modified criterion for logicality will be:

> *Logicality Thesis*: A semantic constraint is logical if and only if it follows from the semantic theory for the language and it is invariant under isomorphisms.

For a semantic constraint to follow from the semantic theory, it must be a true generalization on models according to the theory. A semantic theory for natural language thus, *inter alia*, delineates the logic of natural language. What should be properly included in such a theory will be the subject of the next section. In the meantime, let us test the criterion on our example constraints. If any one of the examples above does not follow from the semantic theory for the relevant language, it will not be considered as logical.[10] But basically, the strengthening of the invariance criterion by the addition of a further condition does not make a difference in the present discussion: we were in any case looking at the semantic theory for natural language. So the last batch of examples (p. 10) would not even be considered, given that none of them would follow from semantic theory. Whether any of the examples in the first batch on p. 10 follows from semantic theory is less clear, and would depend on a more particular specification of methods and domain of inquiry. We leave this specification out of the present discussion.

Now let us thus evaluate the example constraints we drew from Glanzberg:

(a) $I(\text{Ann}) = \text{Ann}$
(b) $I(\text{smokes}) = \lambda x \in D.\ x$ smokes
(c) $I(\text{most}) = \{\langle A, B\rangle \in \mathcal{P}(D)^2 : |A \cap B| > |A \backslash B|\}$

[10] Note that the reference to the meaning of a connective or to a semantic theory only makes sense if we assume that the expressions we deal with are meaningful, and that the language from which they are drawn has an independent semantic theory. Otherwise the aforementioned problem with the invariance criterion does not arise.

Constraint (a) is not invariant under isomorphisms; (b) is invariant only if 'smokes' is trivial—if everything smokes or if nothing does; (c) is invariant under isomorphisms.

We can make a first step here in a characterization of logic in natural language, which I shall defend in the next section. Natural language semantics, formulated model-theoretically can help us model logical consequence in natural language. The criterion of invariance under isomorphisms tells us which semantic clauses provided by natural language semantics characterize its logic. It seems that if there is a logical consequence relation in natural language to be modeled, (c) would be included in the model and (a) and (b) wouldn't. Nevertheless, in the next section I argue that natural language semantics, as a whole, is a model for logical consequence in natural language, even though it appears to include non-invariant constraints. I do this by employing Glanzberg's idea of partiality in semantic theory: while (a) and (b) may both hold, they include elements external to proper natural language semantics. When we weed those elements out, I argue, we are left only with constraints that are invariant under isomorphisms.

3 Semantic theory and logicality

3.1 Glanzberg on partiality in semantic theory

Observe the following clause:

(a) $[\![\text{Ann}]\!] = \text{Ann}$

What does it teach us? That 'Ann' refers to Ann. Assuming competence in the metalanguage (English), and so with the use of 'Ann' in the metalanguage, we can thus tell what 'Ann' in the object language (again, English) refers to. But then, this is no thanks to the given semantic clause. The clause does not appear to be theoretically useful.[11] Similarly, from:
(b) $[\![\text{smokes}]\!] = \lambda x \in D_e.\ x$ smokes
we learn that 'smokes' denotes a function which gives True to whatever smokes and False otherwise. But then, if we are able to successfully apply the function, and so are competent with 'smokes' as used in the metalanguage, it is not this semantic clause that tells us which function it is.
The apparent uninformativeness of such clauses provided by natural language semantics calls for explanation of their value and utility. We first note that these clauses are not *trivial*. They partially determine the role of the words 'Ann' and 'smokes' in English—they say that the former is singular term denoting an individual, and the latter is a predicate of individuals (Higginbotham, 1988, p. 42).

[11] We refer here to semantic theories where the object language is included in the metalanguage, as is customary in contemporary semantic theories for natural language.

By contrast, the following clause seems to deliver significant information on the word 'most':

(c) $[\![\text{most}]\!]_M = \{\langle A, B\rangle \subseteq \mathcal{P}(M)^2 : |A \cap B| > |A \backslash B|\}$

No familiarity with the word 'most' is assumed here. Thus, while clauses (a) and (b) contain a disquotational component, (c) is disquotation-free. Further, we note that 'most', as analysed here, is invariant under isomorphisms. So model-theoretic tools are sufficient and appropriate for giving its meaning in full.

Glanzberg explains that the more a semantic clause leans on disquotation, the less explanatory it is:

> Good explanations tend to appear where we apply model theory or other branches of mathematics to semantics, while mere disquotation signals explanatory weakness (Glanzberg, 2014, p. 268).

The study of determiners through theories of generalized quantifiers has proved fruitful, and indeed explanatory, because of the successful application of mathematical tools in lieu of disquotation. Now, both kinds of clauses, those that lean heavily on disquotation, and those that don't lean on it at all, give us the full meaning of the relevant expressions, in the sense that they give a determinate semantic value. But the former group is partial in the explanatory force of the semantic content attributed. Glanzberg entertains the hypothesis that this partiality will not be overcome by future semantic theories, and so there are aspects of meaning that defy explanation by semantic theory (Glanzberg, 2014, p. 279).

It would be reasonable to associate a discipline with the domain in which its theories are explanatory. In our present case, the theory at stake is semantic theory of natural language, and its domain is linguistic competence. When demarcating the domain through explanatory force, we see that semantic theory is only a partial theory of content, and that where disquotation is appealed to, other theories are needed. Linguistic competence, then, does not provide us with fully determined content: we need other cognitive faculties to achieve that. This, in a nutshell, is Glanzberg's idea. "[S]emantics, narrowly construed as part of our linguistic competence, is only a partial determinant of content" (Glanzberg, 2014, p. 259). Semantic theories have weak explanatory value when it comes to certain expressions. Some lexical items, for example predicates, are provided a type by the semantic theory, and while they are also provided with an extension as their semantic value—what their extension is is not explained by the theory—but is rather taken for granted. This would mean that full truth conditions exceed semantic competence [by contrast to Glanzberg's assumptions in Glanzberg (2015)].

Let us take up Glanzberg's idea, and see what it implies with respect to our main topic, of logicality in natural language. Before we move on, we should qualify the opening statement of the previous paragraph. It is not at all certain that being explanatory is the primary virtue by which theories should be measured. But in our present case, lack of explanation is bound together with lack of informativeness, since disquotational facts lean heavy on prior knowledge of meaning of the same expressions as included in the metalanguage.

If we follow Glanzberg's line, this is where we stand. The way our semantic theories are formulated, they include clauses which fully determine the semantic values of expressions. In some cases, notably of lexical categories, a component of disquotation is included. In other cases, where expressions are more amenable to mathematical treatment, notably in the case of isomorphism invariant expressions, the component of disquotation is null.

Nonetheless, we shall not derive a dichotomous picture of natural language expressions. Glanzberg shows that there are intermediate levels of disquotation and thus of explanatory value. Some semantic clauses, while using some form of disquotation, provide more structure than mere grammatical category. A case in point is that of gradable adjectives. Glanzberg adopts a common approach, by which the semantic value of a gradable adjective is a function from individuals to degrees on a scale.[12] So the meaning of 'tall', for example, is given by:

$$[\![\text{tall}]\!](x) = d$$

Above we have a schema, where d is the degree assigned to x on the scale associated with 'tall'. Adjectival scales in fact have three crucial parameters, each of which must be specified in the lexical entry of any particular gradable adjective: a set of degrees, which represent measurement values; a dimension, which indicates the kind of measurement (height, cost, temperature, speed, volume, and so forth); and an ordering relation on the degrees. The set of degrees will have mathematical properties such as being open or closed on either end. A scale closed on both ends will have minimum and maximum degrees (as in 'full'), a scale open on both ends will have neither (as in 'tall'), and there are scales that are closed on one end and open on the other (as in 'pure' and 'bent').

The dimension, analysed from our perspective, includes the disquotational component. The dimension for 'tall' would be tallness, or height. That of 'flexible' would be bendability. We might or might not use a (near) homophonic translation to the metalanguage, but in any case we use a nearby concept the quality of which provides an interpretation for the degrees on the scale (see Glanzberg, 2014, p. 276). The disquotation is not pure, as it is complemented by the mathematical analysis of the scale.

The ordering relation is a significant additional parameter. The adjectives 'tall' and 'short' share a set of degrees and a dimension, but their ordering relations are inverses of each other. In general, the relation between antonyms can be analysed in this exact manner. Let us take, for example, the antonyms 'wet' and 'dry'. For 'wet', we will have a scale with a minimum (something can be minimally wet—have no water on it at all). For something to be wet, it has to have degree of wetness that is above the minimum. For 'dry' we shall have the same dimension, but an inverted

[12] We rely on Glanzberg's presentation and on Kennedy and McNally (2005). For further references, see Glanzberg (2014, p. 273).

scale which includes a maximum. For something to be dry, it has to be minimally wet. And so we have the entailment (see Glanzberg, 2014, p. 274):

The door is not wet.

entails

The door is dry.

Semantic theory is partially explanatory of the semantic values of gradable adjectives. In usual cases, it will not give a full explication of the dimension associated with the adjective, and in many cases we will have straightforward disquotation. However, we are also provided with some structure that explains entailments as the one above.

The explanatory part of natural language semantics is thus not confined to expressions whose semantic value is invariant under isomorphisms. A gradable adjective such as ‘tall’, while not invariant under isomorphisms, still has an interesting mathematical structure associated with it. And so, we have expressions for which semantic theory gives us a full meaning without appeal to disquotation, expressions where some structure is identified, but the full meaning is given with some appeal to disquotation, and expressions whose semantic clauses merely provide semantic categories, but are otherwise disquotational. In the next subsection we separate out the disquotational element from the relevant semantic clauses. What we shall receive would then be “semantic theory proper” by Glanzberg’s lights. We shall see that if formulated in the framework of semantic constraints, the outcome will satisfy the criterion of logicality we have set in Sect. 2.2.

3.2 Identifying logic in natural language

Let us return to the main theme of this paper. The logic of natural language, I propose, is its isomorphism-invariant part. Further, I conjecture that if we follow Glanzberg demarcation of the explanatory part of semantic theory, we shall see that all of it is invariant under isomorphisms and thus semantic theory can be adequately described as the study of the logic of natural language.

The first step would be to consider semantic clauses as semantic constraints. Recall that we have done so in Sect. 2.1 and obtained:

(a) $I(\text{Ann}) = \text{Ann}$

(b) $I(\text{smokes}) = \lambda x \in D.\, x \text{ smokes}$

(c) $I(\text{most}) = \{\langle A, B\rangle \in \mathcal{P}(D)^2 : |A \cap B| > |A \setminus B|\}$

We can add to these the semantic constraints for ‘tall’ and ‘short’, after appropriately extending the class of possible values of the interpretation function to include degrees, with an order $<$ for height:[13]

[13] We have taken a simplified version of the semantics of ‘tall’ and ‘short’, following Kennedy and McNally (2005). As they point out, the values of ‘tall’ and ‘short’ may be analysed as intervals rather than points (Kennedy & McNally, 2005, n. 17). This, besides complicating the presentation, would not affect our overall argument.

(d) $I(\text{tall})(x) = d$ where the degree of height of x is d.
(e) $I(\text{taller})(x)(y) = T$ iff $I(\text{tall})(x)) > I(\text{tall})(y)$
(f) $I(\text{short})(x) = d$ where the degree of height of x is d.
(g) $I(\text{shorter})(x)(y) = T$ iff $I(\text{short})(x)) < I(\text{short})(y)$

Now, of the semantic constraints above, only the one pertaining to 'most' is invariant under isomorphisms.[14] This is completely in line with our observations in Sect. 2.2. All the constraints above completely fix the meanings of the terms they involve, and so they are invariant under isomorphisms if and only if the term fixed is invariant under isomorphisms.[15]

According to the modified criterion for logicality proposed in Sect. 2.2, a semantic constraint is logical if it follows from the semantic theory for the language and is invariant under isomorphisms. If we take the the semantic constraints above as part of a semantic theory for English, we can look into those following from them that are invariant under isomorphisms, and in this way identify the logic of English. But we were opting for a stronger claim: that *all* of natural language semantics describes (or models) the logic of natural language. This claim is based on Glanzberg's idea of partiality in semantic theory. The semantic constraints above assign a determinate semantic value to the expressions they involve, but by this they outstrip proper semantic theory and its explanatory power.

The natural move at this point is to separate the explanatory from the nonexplanatory in those constraints. Basically, this means that we formulate weaker constraints that do not necessarily give a determinate semantic value, but include no disquotational element. We note that if we follow Glanzberg's line of thought, we have a fundamental tension in semantic theory. On the one hand, the aim of semantic theory is to give (full) truth conditions for sentences compositionally. This means that we need semantic clauses that provide the (full) contribution of expressions to truth conditions, and therefore a determinate semantic value is required. On the other hand, many semantic clauses use disquotation, which is a sign of explanatory weakness. If, indeed, semantic theory is only partly explanatory of content, and we demarcate this discipline according to where it is explanatory, then giving (full) truth conditions goes beyond proper semantic theory. Henceforth we shall accept this conclusion.

The framework of semantic constraints invites us to restrict models vis-à-vis the interpretations of some terms without completely fixing their meaning. Observe the following semantic constraints, which are weakenings of those formulated above:

(a′) $I(\text{Ann}) \in D$
(b′) $I(\text{smokes}) \in \{f : f : D \to \{T, F\}\}$

[14] We adjust the definition of isomorphic models to accommodate degrees as semantic value. As we did with the truth values, we demand that isomorphisms are constant on degrees.

[15] This follows from the proposition mentioned on p. 9, by which a term is invariant under isomorphisms if and only if the semantic constraint completely fixing its meaning is invariant under isomorphisms. See Sagi (2022a).

(c′) $I(\text{most}) = \{\langle A, B\rangle \in \mathcal{P}(D)^2 : |A \cap B| > |A \backslash B|\}$

(d′) $I(\text{tall}) \in \{f : f : D \to S,\ S$ a scale with no minimum or maximum$\}$

(e′) $I(\text{short}) = I(\text{tall})$[16]

(f′) $I(\text{taller})(x)(y) = T$ iff $I(\text{tall})(x)) \succ I(\text{tall})(y)$, where $\succ$ is the order relation of associated with the range for ‘tall’

(g′) $I(\text{shorter})(x)(y) = T$ iff $I(\text{short})(y)) \succ I(\text{short})(x)$, where $\succ$ is the order relation of associated with the range for ‘tall’

These semantic constraints are all invariant under isomorphisms. Also, they include no element of disquotation. Further, it seems that they capture precisely what we have identified as the explanatory part of the semantic clauses from which they originate. In each case, we replaced the disquotational part with a set-theoretic classification of the relevant value: we replaced ‘Ann’ with being a member of the domain; ‘smokes’ with being a function from members of the domain to truth values; ‘most’ was left untouched; the height dimension for ‘tall’ was replaced by a specification of the mathematical properties of the scale; likewise with ‘short’, while the mathematical relation with ‘tall’ remains untouched.

In this way we lose absolute semantic values, and we obtain a range of models. Glanzberg (2015) describes absolute semantics as the aim of natural language semantics. However, there are questions that semantic theory does not aim to solve regarding the extension of expressions (is John in the extension of ‘bachelor’? Is John taller than Ann?). Hence, we have a range of models. Zimmermann (1999) characterizes this range of models as displaying the linguist’s ignorance. Here, we contend that since these questions are beyond the domain of linguistic inquiry, it is not a matter of ignorance but rather of irrelevance.

The expressions we picked are representative of wider classes of lexical categories. The more sophisticated and developed semantic theory gets, the more we can add to the set of semantic constraints forming its logic. The logical consequence relation we shall obtain, if we include a further constraint for ‘John’ and some constraints for the composition of terms, will capture:

Ann is taller than John.

entails

John is shorter than Ann.

[16] The simplified semantics we have attributed to ‘tall’ and ‘short’ ignores an apparent asymmetry, as in:

(a) Tao is seven feet tall.

(b) ??Julian is three feet short.

(Kennedy & McNally, 2005, n. 17). Using intervals as semantic values for ‘tall’ and ‘short’, requiring slightly more complicated constraints, would account for the asymmetry without affecting the overall argument. See n. 13.

We have not fixed ‘taller’ or ‘shorter’ as logical terms—we have only restricted their interpretations by isomorphism invariant semantic constraints.

Presumably, we can formulate isomorphism invariant semantic constraints that will capture the examples discussed in Sect. 1 regarding ‘load’ and ‘cut’. Admittedly, however, these examples pose an additional challenge. As Glanzberg explains, there are good explanations of these kinds of entailments that are not mathematical, but rather have to do with verb categorization or more generally, with “elucidations of meaning” (Higginbotham, 1989).[17] Glanzberg leaves room for the possibility that later versions of these analyses will become more mathematical, but concedes that at this point non-mathematical explanations remain a loose end for his theory. It would be interesting to take Glanzberg’s project further into this realm. In the meantime, we may note that any entailment can be captured by means of isomorphism-invariant semantic constraints. The explanatory value of some such constraints might be doubtful, but if semantic theory predicts an entailment, we can be sure that the relevant semantic constraint will follow from the semantic theory for the language, and the relevant entailment will be included in logic of the language as we characterize it here.

Here is our rejoinder to Glanzberg (2015): there is a logical consequence relation in natural language. While it may be treated as permissive by Glanzberg, we emphasize that it abides by the strictures of invariance under isomorphisms. Once the strict division of terms into logical and nonlogical, completely fixed and variable, is given up, one can accept all the entailments provided by natural language semantics as logical.

Contemporary semantic theory thus models the phenomenon of logic in natural language.[18] Surely, the logic we get via this theory is complicated, always partial and up for refinement and revision. It is nothing like a foundation as envisioned by Frege or the useful framework for science as envisioned by Carnap. Those traditional aims would better find their fulfillment out and away from natural language, as in the work carried out by thinkers in the past 150 years. Yet, if what we look for is an entailment relation where our generalized and modified criterion for logicality is satisfied—that is defined by a set of semantic constraints that are invariant under isomorphisms—then formal semanticists give us just that.

4 Conclusion

This paper included two main claims. First, I presented an alternative framework for logical systems, and claimed that we can formulate an appropriate criterion for logicality—for semantic constraints rather than for logical terms—that can be applied

[17] For verb categorization and analysis of ‘cut’ and ‘load’ and many other examples, see e.g. Dowty (1979), Levin and Hovav (2005).

[18] We have only dealt with examples in an extensional framework, but, as mentioned, the framework of semantic constraints can be extended to an intensional setting, to accommodate the bulk of semantic theory.

in the study of logical consequence in natural language. The second main claim is a strengthening of the first. I used Glanzberg's theory of explanation and partiality in semantic theory to claim that natural language semantics is actually a theory of logic in natural language. Many semanticists claim that their object of study is the logic of natural language—here we give substance to this claim using a notion of logicality accepted by philosophers. Admittedly, the second claim relies on heavy assumptions regarding the study of natural language,[19] and the first on a non-standard conception of logic. The logic we receive might be considered by Glanzberg's position to be a permissive one. Note, however, that this non-standard conception of logic falls right out of the model-theoretic tradition by giving up the assumption that there must be a sharp division of the vocabulary into logical and nonlogical. Holding on to this assumption is a more committed position, even if usually taken for granted. Natural language semantics gives us the opportunity to revisit and let go of this traditional assumption.

Acknoledgements Versions of this paper were presented at the *Logic Colloquium* in Udine, the Philosophy Department in Leeds, the *Anti-Exceptionalism and Pluralisms: from Logics to Mathematics Joint IUSS - Bergen - COGITO Conference* in Pavia and the *Invariance and Objectivity Workshop* in Vienna—I thank the audiences there for comments and discussion. I am also grateful to Gennaro Chierchia, David Kashtan, Ran Lanzet, Simone Picenni, Sebastian Speitel and two referees for this journal for invaluable input.

Author contributions The sole author (Gil Sagi) conducted all the research involved in this work.

Funding Open access funding provided by University of Haifa. This research was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 518/21).

Data availability This work proceeds within a theoretical and mathematical approach, so no datasets are analysed or generated.

Declarations

Conflict of interest The research submitted involves no competing interests.

Ethics approval This work proceeds within a theoretical and mathematical approach, and no ethics approval is relevant.

[19] One can apply the framework of semantic constraints with its invariance criterion to other theories of linguistic competence than what we consider here. For example, Marconi (1997) distinguishes between inferential and referential competence. The former is a competence with entailments, and the latter with application of words to entities to which they refer. In this framework, the logic of natural language would be associated with inferential competence. Presumably, referential competence, for Marconi a part of linguistic competence, would belong for Glanzberg to other cognitive faculties.

References

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention Logic in grammar: Polarity, free choice, and intervention* (Vol. 2). Oxford: Oxford University Press.
Chierchia, G. (2021). On being trivial: Grammar vs. logic. In G. Sagi & J. Woods (Eds.), *The semantic conception of logic* (pp. 227–248). Cambridge: Cambridge University Press.
Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar meaning and grammar*. Cambridge, MA: MIT Press.
Del Pinal, G. (2019). The logicality of language: A new take on triviality, "ungrammaticality'', and logical form. *Noûs, 534*, 785–818.
Dowty, D. R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel.
Gajewski, J. (2002). On analyticity in natural language. Manuscript, https://jon-gajewski.uconn.edu/wp-content/uploads/sites/1784/2016/08/analytic.pdf
Glanzberg, M. (2014). *Explanation and partiality in semantic theory* (pp. 259–292). Metasemantics: New Essays on the Foundations of Meaning.
Glanzberg, M. (2015). Logical consequence and natural language. In C. Caret & O. Hjortland (Eds.), *Foundations of logical consequence*. Oxford: Oxford University Press.
Gómez-Torrente, M. (2002). The problem of logical constants. *The Bulletin of Symbolic Logic, 8*(1), 1–37.
Hanson, W. H. (1997). The concept of logical consequence. *The Philosophical Review, 106*(3), 365–409.
Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar semantics in generative grammar*. Malden, MA: Blackwell.
Higginbotham, J. (1988). Contexts, models, and meanings: A note on the data of semantics. In R. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 29–48). Cambridge: Cambridge University Press.
Higginbotham, J. (1989). Elucidations of meaning. *Linguistics and Philosophy, 12*(4), 465–517.
Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language, 81*(2), 345–381.
Levin, B., & Hovav, M. R. (2005). *Argument realization argument realization*. Cambridge University Press.
Marconi, D. (1997). *Lexical competence lexical competence*. Cambridge, MA: MIT press.
McCarthy, T. (1981). The idea of a logical constant. *The Journal of Philosophy, 78*(9), 499–523.
McGee, V. (1996). Logical operations. *Journal of Philosophical Logic, 25*, 567–580.
Sagi, G. (2014). Formality in logic: From logical terms to semantic constraints. *Logique et Analyse, 57*(227), 259–276.
Sagi, G. (2015). The modal and epistemic arguments against the invariance criterion for logical terms. *Journal of Philosophy, 112*(3), 159–167.
Sagi, G. (2022a). Invariance criteria as meta-constraints. *Bulletin of Symbolic Logic, 28*(1), 104–132.
Sagi, G. (2022b) Considerations on logical consequence and natural language *Dialectica, 999*(1). https://doi.org/10.48106/dial.v74.i2.06
Sher, G. (1996). Did Tarski commit 'Tarski's fallacy'? *The Journal of Symbolic Logic, 61*(2), 653–686.
Woods, J. (2017). Characterizing invariance. *Ergo, 3*, 778–807.
Zimmermann, E. (1999). Meaning postulates and the model-theoretic approach to natural language semantics. *Linguistics and Philosophy, 22*(5), 529–561.
Zimmermann, T. E. (2011). Model-theoretic semantics. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning*. Berlin: Walter de Gruyter.
Zinke, A. (2018). A bullet for invariance: Another argument against the invariance criterion for logical terms. *The Journal of Philosophy, 115*(7), 382–388.