

CHARACTERIZING INVARIANCE

JACK WOODS

University of Leeds

I argue that in order to apply the most common type of criteria for logicality, *invariance criteria*, to natural language, we need to consider both invariance of content—modeled by functions from contexts into extensions—and invariance of character—modeled, *à la* Kaplan, by functions from contexts of use into contents. Logical expressions should be invariant in both senses. If we do not require this, then old objections due to Timothy McCarthy and William Hanson, suitably modified, demonstrate that content invariant expressions can display intuitive marks of non-logicality. If we do require this, we neatly avoid these objections while also managing to demonstrate desirable connections of logicality to necessity. The resulting view is more adequate as a demarcation of the logical expressions of natural language.

1. Introduction

The standard semantic definition of logical consequence, due to Tarski (1936), relies upon selecting in advance a fixed set of logical expressions. We usually do this by enumeration, treating the logical constants in a typical logic textbook as logical and those not not. Since this suffices for applications, many rest content here. However, when we turn to accurately characterizing the concepts of logical truth and logical consequence as they appear in the informal background language with which we actually do mathematics, things get more complicated. Philosophers and logicians have thus attempted to give a philosophically adequate account of why ‘and’ and ‘all’ are logical, why ‘denounced’ and ‘Deniz’ are not, and that settles in a reasonable way disputed cases like = and ‘uncountably many’. The most widespread such account takes logical expressions to be those whose meaning does not depend on the characteristics of particular objects.

The best precise development of this account of logical expressions identifies them with expressions whose extensions are always *invariant* under some class of transformation such as permutations (Tarski 1986), isomorphisms (Sher 1991), strong homomorphisms (Feferman 1999), potential isomorphisms (Bonnay 2008),

Contact: Jack Woods <j.e.woods@gmail.com>

etc. The extension of an expression ϕ is *invariant*, under some transformation T , just in case applying T to the extension of ϕ in a context does not “change”—in some sense—the extension of ϕ . This type of approach, dating back to Tarski (1986) and Mautner (1946), and in some sense, to Kant’s view of logic as an abstraction from the content of judgments (MacFarlane 2002; Tolley 2013; Stang 2014), has a number of useful connections with definability (McGee 1996), and seems to capture at least one sense in which the meaning of logical expressions are immune to the particular features of the objects they are applied to.

Though this approach is commonly applied to formal languages, it is subject to intuitive counterexamples, due to William Hanson (1997) and Timothy McCarthy (1981), when it is attempted as an analysis for natural language expressions. Their examples show that the meaning of a natural language expressions can stay constant while its T -invariant content intuitively varies, in a sense, from context to context. Though there are problems with their examples, exposed recently by Gil Sagi (2015), we can reconstruct versions of their examples which *do* work to demonstrate their point. In light of this, I argue that applying invariance criteria to expressions of natural language, as opposed to formal languages, requires distinguishing two ways in which the meaning of expressions can be immune to the particular nature of the objects they are applied to. Doing so not only solves a number of problems, such as the reconstructed versions of McCarthy and Hanson’s examples, but also highlights important connections between logicity and necessity.

1.1. Invariance Criteria

To see how these approaches to logicity work, consider Tarski’s preferred notion of invariance: permutation invariance. Simplifying slightly, we can model the relevant notion of context, for this limited purpose, as a set of objects (commonly called *a domain*), and then take the extension of expressions on a domain D to correspond to set-theoretic constructions from D —unary predicates corresponding to subsets of D , binary relations to subsets of $D \times D$, unary quantifiers to subsets of the powerset of D , etc. The *general extension* of an expression is a function from the set of domains to set-theoretic constructions of the appropriate type over these domains. For example, the general extension of ‘is a cat’ is the function whose value for any domain D is the set of cats in D . In the actual context—ours—the extension of ‘is a cat’ is simply the set of cats. The general extension of ‘for all’ is the function whose value for a domain D is $\{D\}$. The general extension of ‘there is a’ is the function whose value for a domain D is the set of non-empty subsets of D .

Given a permutation π , we define π^+ , a function on set-theoretic constructions out of D , as follows: $\pi^+(e) = \{\pi^+(d) \mid d \in e\}$ and $\pi^+(e) = \pi(e)$ if $e \in D$. Then we can say that the general extension of ϕ is permutation invariant just

in case, for every permutation π and every domain D , ϕ^D (the extension of ϕ on D) does not change when we apply π^+ to it—i.e., when $\pi^+(\phi^D) = \phi^D$. So, the general extension of the universal quantifier ‘for all’—which can be modeled by a function from domains to their singletons—is permutation invariant since permuting D results in D , (i.e., $\pi^+(\{D\}) = \{D\}$). The general extension of the predicate ‘is the greatest Polish logician’, on the other hand, is not permutation invariant as there is a permutation π on a Tarski-containing domain D such that $\pi(\text{Tarski}) = \text{Starsky}$. Given this permutation, $\pi^+(\text{‘is the greatest Polish logician’}^D) = \{\pi(\text{Tarski})\} = \{\text{Starsky}\}$.¹ A little work establishes that all truth-functions, the existential and universal quantifier, the identity relation and the non-identity relation, and cardinality quantifiers are permutation invariant.

Although invariance-type accounts of logicality are common, their justification as useful indicators of logicality is slightly difficult to pin down. The rough background thought for the case of permutation invariance is put well by Vann McGee:

Any operation which is disturbed by a permutation must somehow discriminate among individuals in the domain, and any consideration which discriminates among individuals lies beyond the reach of logic, whose concerns are entirely general. (1996: 567)

McGee’s idea, I take it, is that the *formal* notion of invariance measures, somehow, how much the general extension of an expression carries information about individuals. Truly logical expressions, such as those found in the usual logic books, have extensions which do not change when we swap around the facts concerning individuals in the context in which they are applied. The extensions of expressions like ‘cat’, ‘sat’, and ‘Matt’, however, can change rather dramatically. This is supposed to show that the “meanings” of non-logical expressions, in at least one sense of meaning, encodes information about the underlying types of objects they apply to, whereas the “meaning” of logical expressions encode no such information.

Formal criteria for logicality, such as permutation invariance, are really criteria that apply to model-theoretic *surrogates* for natural language meanings. Sometimes, we can more or less identify the model-theoretic surrogates for natural language meanings with their actual meanings. Perhaps this is true for mathematical language; perhaps it is true of a reformed language of science like that dreamed of by both Frege and the logical positivists. Sometimes, we need to carefully distinguish mathematical models of meaning from meanings themselves, such as when we deal with natural language in all of its modal, aspectual, and colored character. However, either way, formal criteria for logicality inform our choices in modeling natural meaning by telling us which expressions we ought to

1. For details and extensive further discussion, see Woods (2014).

treat as logical expressions in our models. In particular, it tells us that we should treat as logical those expressions whose meaning, as modeled by our formal models, is invariant—i.e., those expressions whose general extensions are invariant. It is also important to note that invariance criteria are *immanent*, not *transcendent* criteria—they presume that the metatheoretic resources with which we describe the semantic features of expressions are already understood. It would be desirable to avoid this property, but it is entirely unclear how to do so.²

The formal criteria tells us to treat these expressions as logical because their meaning, as modeled by the model-theoretic surrogates, displays a property which we take to represent the intuitive property we assign to the natural language meaning of these expressions. Transformations model ways in which the underlying context could be different: permutations model how we could swap around the facts concerning the objects in a domain, isomorphisms model how we could swap around which objects have which features and, additionally, which objects there are, so long as there are the same (cardinal) number of them, etc. *T-invariance*, then, for some class of transformation *T*, represents a property of the meaning of certain expressions—that their meaning is not dependent on particular features of the objects in the domain. So long as we have successfully modeled the meaning of the relevant expressions and so long as *T-invariance* really does represent the intuitive property of non-dependence on particular features, we are licensed in claiming that the actual meaning of these expressions really does display this intuitive property of logicity.

Inspection of the justification just sketched, though, suggests that using permutation invariance as the relevant transformation is simply too narrow. There are definable permutation-invariant general extensions which correspond to intuitively non-logical meanings, so permutation invariance is not the most discriminating criterion we could apply.³ Consider, for example, the wombat quantifier, whose extension on domains with wombats is the same as that of the universal quantifier, but the existential quantifier on domains suffering from wombatlessness (McGee 1996). Even though the extension of this operator on each domain is permutation invariant, which extension it has depends quite heavily on non-logical information about the domain, such as whether it contains wombats. In order to screen these types of operators out and to have a more discriminating criterion, theorists have generally opted for transformations *between* domains, such as isomorphisms.⁴ It is easy to show that even though McGee's funky quantifier

2. I do not have space here to discuss the relationship between semantic criteria and proof-theoretic criteria for logicity. My aim here is to make more plausible the best going semantic criteria, not to argue against entirely alternative approaches.

3. Although Woods (2014) and Dutilh Novaes (2014) further argue that permutation invariance is not even necessary for logicity. The former argues that the criterion can be patched. The latter does not.

4. In this context, an isomorphism between domains is simply a bijection. The reader is

is permutation invariant, it is not isomorphism invariant.

Note that since extensions which fail to be permutation invariant *eo ipso* fail to be isomorphism invariant, McGee's example does not show that permutation invariance is not a necessary condition. It merely shows that it is insufficiently discriminating to be the most useful guide to logicity. More could and has been done to narrow the class of logical expressions by choice of transformations, as work by Feferman (1999) and Bonnay (2008) testifies to. Perhaps isomorphism-invariance is not the best choice of transformation for this approach to logicity. Nevertheless, for simplicity, we will focus on isomorphism invariance in what follows as my overall point does not turn on which type of transformation between domains we choose.⁵ I will flag below where the limitations of isomorphism invariance crop up (Sections 3–4), though I will note at the outset that we could relieve some of these limitations, at the cost of significant complexity, by moving to a more discriminating transformation such as Feferman's *strong homomorphisms*. Taking isomorphism invariance as the relevant transformation, we can put the standard invariance criterion as follows:

SIMPLE INVARIANCE An expression ϕ is logical only if its extension is isomorphism invariant.

I will presume that, insofar as it goes, SIMPLE INVARIANCE does capture a type of independence of the meaning of logical expressions from particular features of objects they are applied to. Defending this approach to logicity is beyond the scope of this essay; my primary aim here is to improve it.⁶

1.2. Invariance and Meaning

Of course, the conception of meaning which is modeled by merely by the extension of expressions is impoverished; this has sat badly with a number of people. Surely, we might think, in order for an expression to be logical, its *real meaning*, not merely its *extension*, needs not discriminate among the objects in the

cautioned to keep this unfortunately well-established terminological choice in mind below and when reading the literature on invariance criteria generally.

5. Many theorists, such as Sher (1991), use invariance as only one necessary criterion in a list of jointly sufficient criteria for being a logical expression. I discuss Sher's view extensively below, but it is worth noting at the outset that isomorphism invariance is clearly the principal criterion on the list and has been taken as more-or-less a necessary and sufficient condition (perhaps with the addition of Sher's point about rigidity, described below) by many after her.

6. Recent work by Tim Button and Sean Walsh (in press) suggests that this defense will require more than might initially be thought; in particular, one needs to argue in detail that the sort of non-dependence picked out by isomorphism invariance (or a close cousin) really does capture an intuitive non-dependence property characteristic of the meanings of logical expressions. I remain hopeful that this can be done and hope, in particular, to address their objections and their relation to the work here in another venue.

relevant domain.⁷ This complaint isn't just that extension invariance is insufficient for a full account of logicity for natural language. The real complaint is that extensions are inadequate as an account of the meaning of expressions, even if sometimes the extension of an expression adequately models the natural language meaning.

Timothy McCarthy (1981) and William Hanson (1997) have offered compelling arguments along these lines against the invariance criterion when used as an account of the logicity of expressions (in the way described above).⁸ Their complaint is that we can use non-logical materials to fix the extensions of expressions in a way which intuitively modifies the logical properties of their meaning, but which does not modify the invariance properties of their extensions. Their arguments go roughly as follows:

1. Invariance is a formal property applying to *extensions* of expressions, not their meanings.
2. The extension of an invariant expression can be fixed by non-logical ingredients.
3. Expressions whose extensions are fixed by non-logical means are thereby non-logical.
4. Therefore, invariance of extension is insufficient for the logicity of an expression.

(1) is established by inspection of the definition of (simple) invariance; (2) can be established by looking at semantic rules for certain stipulated operators, defined over contexts as follows:

$$(N) \forall c [N\phi' \text{ holds at } c \leftrightarrow (\phi' \text{ holds at } c \text{ iff } \neg K)] \text{ (McCarthy 1981)}^9$$

where K is some contingent actual truth; (3) is *prima facie* plausible, given that defined operators like N seem to display non-logical behavior (see below), and

7. I am here privileging what Bonnay (2008: 6) calls *the formality argument*. I do not mean to suggest that this is the only way into the invariance criterion; see Bonnay's discussion of *the generality argument* and the relationship between the two of them. The issue of what these joint arguments support is aimed at picking which transformations matter rather than at my question below about whether we should be concerned with invariance of extensions, meanings, or both, so I will bracket it.

8. Few have seriously criticized the invariance criterion in its original Tarskian usage as a way of discriminating a set of logical (set-theoretic) objects from non-logical objects, though see (Button & Walsh in press: Chapter 8). See Sher (2001; 2003) for a useful reminder that we ought distinguish the two projects. I offer some reasons to broaden the Tarskian criterion in Woods (2014), but even these are meant as a modification, not a rejection of Tarski's approach.

9. I have switched McCarthy's talk of structures for talk of contexts throughout to generalize the point and emphasize that little depends on how we represent the intuitive idea.

(4) follows from (1–3). The essential upshot is that invariance criteria fail to capture aspects of the meaning of expressions that plausibly can interfere with their logical character and, hence, are at best a weak *necessary* condition of logicity.

Gil Sagi, in recent work, has criticized these arguments (Sagi 2015). Although Sagi’s criticisms of this argument are instructive, there is nonetheless a way to understand McCarthy’s and Hanson’s arguments—and, in particular, to understand (3)—that avoids Sagi’s objections and which is compelling against versions of the invariance criterion (Sections 2.1, 4). Let me be clear: I accept Sagi’s criticisms against a certain way of taking McCarthy and Hanson’s arguments; so construed, the arguments fail. Yet, viewed another way, we can rescue the point that McCarthy and Hanson are drawing out attention to and, in so doing, give a more plausible account of how to use invariance as a criterion for logicity.¹⁰ I take part of the implicit, if not explicit, point of McCarthy and Hanson’s arguments to be that we need to take a broader view of meaning to obtain an adequate general account of logical expressions for natural language. This has gone relatively unnoticed as philosophers interested in logicity are usually interested in distinguishing between logical and non-logical expressions in languages which have an extensional flavor, such as a large fragment of mathematical language.

Below, I use Sagi’s discussion to elucidate the way in which SIMPLE INVARIANCE falls short of a characterization of logicity for natural language expressions; namely, by being defined only over an impoverished account of meaning (Section 2). We can do better than SIMPLE INVARIANCE by adding to invariance of extension—as used by SIMPLE INVARIANCE—a requirement that logical expressions also have invariant meanings (cashed out here in terms of Kaplanian characters, though the point is general.) Roughly, we will say that a Kaplanian character for an expression ϕ is isomorphism invariant when, fixing its content at a context of use c , applying any *index-preserving* isomorphism i from c to d results in the content which would have been fixed for ϕ at d (further technical details are put off until the appendix; the intuitive idea should suffice for now.) Both types of invariance are important; each represents different aspects of how logical expressions are insensitive to the underlying natures of objects.

1.3. Broadening Meaning, Broadening Invariance

We start by putting away an initial suggestion. Moving to a slightly more encompassing picture of content, such as *intensions* (i.e., functions from contexts to particular extensions), will not solve this problem. Inspection of the above example from McCarthy will show that it can be taken as laying down a function from contexts to truth-functions. If that content is itself invariant, in the relevant

10. I believe that Hanson’s argument, as opposed to McCarthy’s, is best interpreted in the way that I will suggest. However, as my point is not to explicate McCarthy or Hanson’s prose, I shall leave interpretive issues at that.

sense, then McCarthy and Hanson's point stands since N is still intuitively non-logical. So we need an even more encompassing notion than that of an intension to capture the point that McCarthy and Hanson are making. In fact, since we will deal here with languages which do not contain modal expressions and since we are taking the relevant form of transformation to be isomorphism invariance, we can avoid worrying overmuch about the difference between intensions and extensions, taking the intension of an expression to be invariant when its extension, in every context, is invariant (i.e., identifying intensions with what I called above 'general extensions.')

For unmodalized languages, this suffices.¹¹ So I will henceforth talk in terms of contents—that is, general extensions—being invariant and contrast this with meanings being invariant.

The key to addressing McCarthy and Hanson's complaint is to broaden the target notion of meaning in a different direction. Distinguish between contexts of evaluation of a statement—contexts where it is evaluated for truth and falsity—and the context of use of a statement—the context where we assert it, utter it, or what have you.¹² These can come apart: my assertion that I'm writing at the moment is true when evaluated in its context of use, but not true when evaluated at many other contexts, such as the one where I'm sensibly taking a mid-afternoon constitutional.

What is causing the problem in the above example is that invariance of content tracks insensitivity of the meaning of logical expressions to features of objects in contexts of evaluation whereas invariance of *meaning* ought to track insensitivity of logical expressions to both features of objects in contexts of evaluation and in the context of use. I will develop this point below, but for now it suffices to note the distinction between these two roles that context can play.

Now if logical expressions really have extensional meanings, then their meanings won't change even if we apply our transformation to the context of use. But if we are concerned to isolate the class of logical constants within natural language, then we ought not presume the extensional character of logical expressions *at the outset*, but hope that it follows from criteria which are motivated by underlying intuitions about the nature of logic, such as the insensitivity of logic to the particular features of things. Taking content invariance—insensitivity to the nature of things in the *context of evaluation*—as our criterion does not secure this result as examples like McCarthy's demonstrate. I will discuss in more detail below the construction of expressions with non-extensional meanings which are nevertheless content invariant.

11. Dealing with modalized language adds an entire layer of complexity; I hope to return to this complexity in future work.

12. *Nota Bene*: 'contexts of evaluation' are not the same as the modern-day relativist's contexts of assessment; rather, these are what Kaplan calls *circumstances*. Moving to a third role for contexts, assessment in the relativist's sense, opens the possibility for a third type of invariance. I hope to return to this issue elsewhere. Thanks to Julien Murzi for discussion.

Some theorists, such as Sher (1991), avoid this worry by stipulating that we consider only extensions when *introducing* logical expressions for a constructed (extensional) logical language.¹³ Accepting this condition *presupposes* that the resulting expressions will be extensional; part of the implicit point of Hanson and McCarthy is that if we do not presuppose this, then we can define expressions which rigidly denote isomorphism-invariant objects, yet whose intuitive meaning is not extensional. So, even if our task is to lay down conditions on how to introduce logical notions into logical languages, we ought to also impose invariance of meaning as a constraint instead of stipulating that we only consider the extensions when introducing logical expressions. The alternative of presupposing that all logical notions are extensional undermines much of the interest in invariance criteria as an explication of the nature of logical expressions. Sher's stipulation may be adequate for her purposes, but it seems ad hoc if our aim is a general semantic criterion for logical expressions in, and introduced by, natural language. (See Section 2.2).

In brief, I claim that we can make sense both of invariant *extensions* and invariant *meanings*—construed in terms of Kaplanian characters—and that invariance of *both* sorts is required for an expression to be logical in the fullest sense.¹⁴ So, presuming that isomorphisms are the right transformation here, I advocate:

DOUBLE-STANDARD INVARIANCE An expression ϕ is logical only if both the *content* and the *character* of ϕ are isomorphism-invariant.

as a more plausible version of the invariance criterion. And, even if one rejects invariance criteria as a full account of logicity, I advocate DOUBLE-STANDARD INVARIANCE as at least a necessary condition on logicity as it has significant theoretical payoffs:

- DOUBLE-STANDARD INVARIANCE does a better job at isolating the logical constants than SIMPLE INVARIANCE in contexts where we allow that meaning comes apart from general extensions. (Sections 2.1, 2.2)
- DOUBLE-STANDARD INVARIANCE has desirable connections not only to ordinary necessity, but so-called deep necessity. (Section 3)
- Logical truths involving only character-invariant expressions, such as those meeting DOUBLE-STANDARD INVARIANCE, are plausibly *a priori* graspable as true from a grasp of the meaning of the constitutive expressions, explicating the sense in which logical truths are *a priori* (subject to caveats discussed below). (Section 3)

13. See Sher's (1991: 54–56) discussion of condition (B).

14. I use Kaplanian characters for simplicity, but the approach could be adapted to different semantic pictures. Bonnay, in his dissertation, suggests a similar approach using a Stalnakerian approach to content; I hope to discuss the relation between these two approaches elsewhere.

- DOUBLE-STANDARD INVARIANCE explains and entails a desirable *rigidity* constraint that other theorists have imposed on logical expressions. (Section 4)
- The two standards making up DOUBLE-STANDARD INVARIANCE, content and character invariance, are unified in the sense that both represent types of insensitivity to particular objects. (Section 5)

My aim is not criticism of any particular theorist's version of the invariance criterion; as will become clear below, theorists have been somewhat sensitive to points in this area (Sher 2001). However, DOUBLE-STANDARD INVARIANCE *explains* certain restrictions we might place on logical expressions in terms of something more fundamental such as character invariance. My methodology is thus to first use Sagi's discussion to elucidate the need for character invariance and argue for the appropriateness of considering both types of meaning when investigating the logical properties of expressions (Sections 2–2.2). I then turn to connecting up character and content invariance with the modal and epistemic properties that Hanson and McCarthy's arguments suggest logical expressions ought to have (Section 3), and use the results to reformulate their argument in a way immune from Sagi's criticisms. In Section 4, I discuss reasons to prefer DOUBLE-STANDARD INVARIANCE over the combination of something like SIMPLE INVARIANCE and a constraint implying that logical constants are rigid designators. Sher (1991) uses a combination like this for logical languages; my main point in Section 4 is that for natural languages, it is insufficient. I close by drawing together the reasons to prefer DOUBLE-STANDARD INVARIANCE to SIMPLE INVARIANCE and, finally, in a brief appendix, I lay out some technical details of the connections between invariance and necessity.

2. Sagi's Response to McCarthy and Hanson

Sagi distinguishes two distinct versions of McCarthy and Hanson's arguments, the *modal* and the *epistemic*, which focus on different ways to motivate and understand the third premise. On the modal interpretation, it claims that if the extension of an expression is fixed by non-logical materials, then it is possible for logical truths to be false, contra the extremely plausible claim that logical truths are necessary.¹⁵ On the epistemic interpretation, the third premise claims

15. The objection is not that classical "logical" truths like $\exists x x = x$, which are intuitively contingent, come out necessary—this is unavoidable given the assumption that domains are non-empty and, perhaps, a pragmatic justification of this could be given. It is rather that some logical truths will come out as contingent on this way of specifying logical expressions *even if* we hold cardinality facts fixed. For discussion of truths constructed out of logical materials which are equivalent to facts about the size of the domain, see Woods (2014: 302–303).

that if we fix the extension of logical expressions by *a posteriori* materials, then there will be logical truths which we could not know *a priori* since we could not know, in the absence of some *a posteriori* knowledge, that the logical truth is true.

McCarthy's and Hanson's arguments are about the logical properties of expressions in natural language. This is clear from context; if it were simply an abstruse model-theoretic property at issue, arguments involving the intuitive necessity of 'logical' expressions and the *a priori* status of certain truths involving them would be out of place. So when we distinguish between the *metalanguage* with which we define certain constants and the *object language* in which they are defined, we are really distinguishing between two ways of taking our own language or some rigorized version thereof, at least in some salient cases. We are distinguishing between the language with which we introduce an expression (call it *PseudoEnglish*₁) and the resulting language under study in which it has been introduced (*PseudoEnglish*₂). I shall understand McCarthy's and Hanson's examples this way in what follows. One reason for this, suggested above, is that identifying logical expressions only for a rarified collection of formal languages is of significantly less interest than identifying logical expressions in natural language or close analogues thereof.¹⁶ Note, briefly, that we have said little at this point about the details of *PseudoEnglish*₁—it may contain non-extensional expressions or it may not—and the character of legitimate definitions in *PseudoEnglish*₁—these may contain non-extensional expressions or they may not.

Sagi objects to the modal interpretation by pointing out that $N\phi \leftrightarrow \neg\phi$, a statement of *PseudoEnglish*₂, actually is necessarily true. Yes, if *K* were false, the statement expressed by the sentence ' $N\phi \leftrightarrow \neg\phi$ ' would not have been necessarily true, but this would be a case in which the *PseudoEnglish*₁ definition of *N* yielded an expression with a different meaning, resulting in a new language *PseudoEnglish*₃. The *PseudoEnglish*₃ statement $N \leftrightarrow \neg\phi$ is false, but what of it? *PseudoEnglish*₃ is a different language. In brief, Sagi argues, correctly in my view, that if we distinguish the sentence ' $N\phi \leftrightarrow \neg\phi$ ' from the interpreted statement $N\phi \leftrightarrow \neg\phi$, we can see that the former could have been false, but not the latter. But the possible falsity of the former is not of any interest. Against the epistemic interpretation, she again notes that the sentence ' $N\phi \leftrightarrow \neg\phi$ ' can have its meaning determined by semantic rules involving *a posteriori* materials without obviously transferring the *a posteriori* status onto the statement so determined. Sagi refrains from pursuing the epistemic objection on grounds that it would require saying far more about the *a priori*.

I agree with both of Sagi's worries. Her objection to the modal interpretation is obviously cogent and her worry about the transference of *a posteriori* materials from the account of the meaning of an expression to the *a posteriori* status of logical truths involving this expression is apt. We can, however, give an al-

16. For relevant discussion, see Section 4 below.

ternative version of the argument which skirts Sagi's objections. Hanson is best understood as claiming that Sher's account allows that the logical status of the extension of an expression may crosscut the logical status of its *meaning* (Hanson 1991: 393–4). And McCarthy puts the upshot of his argument as follows:

If logical constanhood is a semantic property, the semantic description of a logical constant is not completely determined by the class of functions it introduces; for the suggested examples provide us with semantically distinguishable constants that are associated with the same function. (McCarthy 1981: 516)

So, again, if the relevant sense of meaning for the logical status of expressions is not exhausted by general extensions, then Sagi's defense is incomplete when viewed as a defense of SIMPLE INVARIANCE as a criterion for logicity for expressions of natural language.

2.1. *Rejiggering Hanson and McCarthy's Examples*

Let us introduce some machinery to capture the aspects of meaning that McCarthy and Hanson are concerned with. Following Kaplan (1989), we can treat the *contents* of expressions as functions from contexts of evaluation to extensions. Meaning, on the other hand, we will take to be *character*, where the character of an expression is a semantic rule which can be modeled by a function from context (of use) to contents.¹⁷ The character of the first-person pronoun, for example, is the semantic rule:

‘I’ refers to the speaker (of the context of use)

Some expressions, such as ‘water’, arguably have constant characters—from any context of use, the content of ‘water’ is the function from contexts of evaluation to the water within. Others, such as ‘I’, are not constant, even though the content of the expressions is *rigid* in Kripke's sense—in any context of use, ‘I’ takes us to the constant function from context of evaluation to the actual speaker of the context of use. That is, the *content* of ‘I’, uttered by me, is the constant function which takes every context to me, but the *content* of ‘I’, uttered by you, is the constant function which takes every context to you. It is important to note that speakers can grasp the character of an expression without fully grasping its content. As famously noted by Kaplan (1989) and Perry (1979), we may not know

17. For Kaplan, strictly speaking contents *model* the notion of a *propositional constituent* (the sort of thing figuring in a proposition), whereas *characters* model the notion of a *semantic rule*. For our purposes here, we can occasionally speak with the vulgar and identify what is modeled with what models it. They will be distinguished where it matters. Thanks to Eliot Michaelson for discussion.

which object in particular 'I' refers to while knowing it refers to whoever spoke it. Eavesdropping, I may not know that the content of 'I', for example, refers to Jack while knowing full well it refers to whomever said it. This ignorance of the particulars of the context of use is sufficient to destroy our grasp of the content of 'I' without destroying our grasp of its character.

For many expressions, there is no substantive difference between the character and content of the expressions involved—at least on one plausible picture of how the meaning of intuitively non-logical expressions works (Kaplan 1989). On the rival picture, many expressions have non-constant character, much like indexicals, but this will not often affect the invariance properties of the content since both the character and the content of intuitively non-logical expressions will be variant (Jackson 1998). Either way, when we are dealing with such expressions, we need not worry overmuch about distinguishing between character and content. But when we explicitly define *semantical rules* and view these as meaning-constitutive, as Sagi suggests we interpret Hanson and McCarthy as doing, then we need to worry about the difference between context of use and context of evaluation and, correspondingly, the difference between character and content. In different contexts, we can use an expression defined by the same semantical rule without a difference of meaning—i.e., a difference in character—just a difference in content.

Consider again the sort of example that McCarthy used to make his point. We have a true contingent claim, K , and we lay down the following excruciatingly explicit definition of an expression, 'N', speaking PseudoEnglish₁:

$$(N) \forall c \forall \phi ['N\phi' \text{ holds at } c \leftrightarrow (\phi \text{ holds at } c \text{ iff } \neg K)]$$

resulting in an extended language, PseudoEnglish₂. Now, given that K is true in our context of use—the context of definition—what this says is that in any context of evaluation c , $N\phi$ is satisfied if and only if $\neg\phi$ is. That is, (N) specifies a function from contexts of evaluation to the truth-function \neg (the truth-function taking us from **T** to **F** and conversely); this is the *content* of 'N' in PseudoEnglish₂.

Now, what about its character? Settling this requires distinguishing between two ways of interpreting definitions like (N). We could treat the definition as *conditional*—if K is true, we have defined N as coextensional with \neg ; if K is false, we have defined N as coextensional with the trivial functor τ taking us from **T** to **T** and similarly for **F**.¹⁸ This is to read (N) as laying down an expression with constant character; so read, N does not shift its meaning from context of use to context of use. If this is how we are to read N , then I am entirely in agreement with Sagi that there is no sense to be made of how $N\phi \leftrightarrow \neg\phi$ could have been

18. *Nota Bene*: (N) is not an abbreviative definition in PseudoEnglish₁. Sagi is right to distinguish that case from the case of definition in a metalanguage. It is in this latter sense I intend both interpretations of (N).

contingent given that K is actually true. If $N\phi \leftrightarrow \neg\phi$ is false, it must be that we defined N in a context of use in which K was false. Then it's logically false and so because we speak a different extension of *PseudoEnglish*₁, *PseudoEnglish*₃, in which N has a different meaning. But this conditional interpretation is not the only way to understand this type of definition.

We could alternatively understand (N) as directly specifying in *PseudoEnglish*₁ a semantic rule—the *character* of N —resulting in a context-sensitive operator in *PseudoEnglish*₂.¹⁹ When I say N is context-sensitive, I mean that its *content* can change from context of use to context of use, not that its character or meaning so changes. Rather, like 'I', 'here', and 'now', the meaning of N is given by a particular semantic rule and, given a particular context of use, it has constant content, just as with the first-person indexical. Regardless of how we *should* understand McCarthy's and Hanson's definitions, we *can* explicitly define an expression this way. Given an contingent claim K , I hereby specify the character of (N^i), as a semantic rule thus:

$$(N^i) \forall c [N^i\phi \text{ holds at } c \leftrightarrow (\phi \text{ holds at } c \text{ iff } \neg K)]$$

The *content* of N^i , given that K holds at the actual world of our context, is \neg , but in another context, it could be the trivial functor τ . But it does not follow that N^i has a different meaning in that context insofar as we take meaning to be given by character, not content.

Note that we could make the same point by introducing an indexical term directly. Suppose, starting with a language which did not contain the first-person indexical, we laid down the following definition of it:

$$(I^*) \forall c [I^* \text{ refers, at } c, \text{ to the speaker}]$$

This definition, as with our definition of N above, could be interpreted in at least two importantly distinct ways.²⁰ Read as a conditional definition, it specifies a term with constant character which refers at any context of evaluation to the speaker at the original context of definition. Read as a specification of a semantic rule, it lays down a term with shifting character which, at any context of evaluation, refers to the speaker at the context in which I^* was used. From some contexts of use, I^* , when its definition is interpreted as laying down a context-sensitive semantic rule, has different content than from other contexts

19. (N) does not explicitly make use of any non-extensional materials; that is, abstracting from the quantification over contexts necessary to make clear which expressions are being used and which are being mentioned, none of the content of the semantic rule is non-extensional. We could, of course, use non-extensional materials in laying down the definition, without changing the overall point in the slightest.

20. 'Speaker' could also be interpreted as elliptical for 'speaker at c '. This is not the intended reading, so put it aside.

of use. Speakers from these differing contexts nevertheless speak the same language even though their uses of I^* have different contents, just as you and I are speaking the same language and using expressions with the same meaning when we each utter the sentence ‘I’m talking here.’ As with I^* , so with N^i .

Now, $N^i\phi \leftrightarrow \neg\phi$ is dependent, in a sense, on contingent materials—since K is true in the context of use, $N^i\phi \leftrightarrow \neg\phi$ is necessarily true, but it would not have been true, necessarily or otherwise, had we evaluated the semantic rule for N^i in a different context of use. Yet N^i rigidly refers (in the sense specified by Sher, discussed below); its content, once fixed by a context of use, is the same in every context of evaluation, just like ‘I’ (for more discussion, see below). As long as a significant portion of meaning has to do with such semantic rules, we can make sense of *one way* a claim like $N^i\phi \leftrightarrow \neg\phi$ is contingent, even though it is true no matter which context we evaluate it at.

2.2. *The Relevant Sense of Meaning*

Should we assume, though, that this is the right sense of meaning for logicity? As I pointed out above, it seems to me that we have strong methodological reasons to *start* with a notion of meaning which is adequate to analyzing natural language—if our task is to identify the logical constants of a natural language or even a quasi-natural language like that used in philosophical contexts, then we should not presume at the outset that all such logical expressions will have purely extensional meanings. Rather, we should start with the natural conception of meaning for the language under consideration, especially such language that figures into arguments we might be interested in assessing for validity, and see if explicating intuitions about the properties of logical expressions returns the result that their meaning is extensional.

Viewed this way, it seems inescapable that the relevant sense of meaning encompasses more than extensions. We can, after all, reason from ‘I am here’ to ‘someone is here’ or from ‘There is a tall student in my class’ to ‘There are tall students’. We mis-analyze the meaning of the premises if we treat indexical terms like ‘I’ or gradable adjectives like ‘tall’ as having extensional meanings or even intensional meanings *even when* we could capture the validity of these arguments by so doing. In fact, a wide variety of arguments philosophers are interested in analyzing involve indexical expressions as well as actuality and tense operators, their locative analogues such as ‘here’ and ‘now’, and the like. Why should we restrict our account of logical expressions, before we even start our investigation, to those expressions particularly focused on in the fragment of the initial formalization of logical languages aimed at representing proofs in *mathematics*?

Even in that rarified domain, we may note that there are expressions which can be usefully theorized about by distinguishing between character and content.

ϵ , treated as an arbitrary choice operator, is one nice example.²¹ Plausibly indefinitely extensible notions like ‘definable’ or ‘ordinal’, which can be treated as having different meanings at different stages-of-definition, are another. Taking a broader picture of meaning, even in mathematics, potentially gives us a better sense of *why* certain expressions are well-treated as logical-in-the-fullest sense and why certain expressions which are *like* these in some senses, but not others, are better seen as not being logical in the fullest sense.

A related reason for not identifying the relevant sense of meaning for logicity as well-modeled by general extensions is that doing so occludes different senses of logicity. “Logical truths” of the logic of demonstratives, like ‘I am here now’, exhibit a number of marks of logicity. They are indifferent, in a sense, to the underlying natures of the objects they are applied to, even though they are not isomorphism invariant. Likewise, it seems to me that we could take a permissive stance on operators like N^i and claim that they are also logical in a weak sense—they exhibit some, but not all, of the marks of logicity. When we distinguish invariance of character and invariance of content, we can explain why certain operators are only logical in a relaxed sense—they’re only halfway invariant. Logicity in the fullest sense requires invariance of both content and character since these model two different ways the meaning of an expression could be sensitive to the natures of objects. But doing so requires that we recognize, at the outset, the two senses of meaning on offer.

So it is best to assume that the relevant sense of meaning is not exhausted by extensions or intensions. Summarizing the discussion so far, definitions, like that of McCarthy’s N , can be viewed as determining not only *content*, but also *character*, corresponding to a *semantic rule* specified by the definition. And since we can define—even in an extensional *PseudoEnglish*₁—expressions with characters which characterize distinct, yet invariant *contents* in different contexts of use without changing their meaning, there is a *prima facie* mismatch between the invariance of the *content* of expressions and the logicity of expressions. This mismatch is made more plausible by the contingency—in a sense—of some truths marked logical by the content invariance criterion and our inability to justify such as logical truths by *a priori* methods. Our account of the *meaning* of a logical constant should thus involve both the character of the constant and its content. The upshot is not that we should avoid using the invariance criterion in order to account for logicity, but rather that this criterion should be applied to both the

21. I have argued elsewhere that we should recognize arbitrary choice operators like ϵ as logical operators in more-or-less the fullest sense and shown that it is possible to set up a principled account of the denotations of expressions which does so (Woods 2014). The character/content distinction, however, provides a fallback position for those, like Breckinridge and Magidor (2012) who treat such operators having a standard, though arbitrarily specified, denotation. They can treat the choice of an arbitrary representative for each predicate as part of the context and let ϵ have invariant character and wildly variant content.

character of an expression and its content. Only if both are invariant should we regard the expression as logical.

3. Contingency and Invariance

A secondary aim of McCarthy and Hanson's examples is emphasizing that necessity, in some form, is itself a criterion of logicity. Logical expressions, intuitively and traditionally, should yield truths which are necessary (Hanson 1997). The necessity of logical truths is, presumably, supposed to be a *consequence* of the insensitivity of logical expressions to the natures of objects they are applied to. At least, a criterion of the logicity of expressions which entailed that all truths composed only of logical expressions were necessary would be all the more plausible. This because it would capture one important criterion of logicity on the basis of a plausibly more fundamental criterion of logicity. However, there is a problem here. We now have two senses of meaning—content and character—in play and, correspondingly, two senses of insensitivity in play. Invariance of character, even intuitively understood, does not guarantee necessity in the usual truth-at-all-worlds sense. This can easily be seen by noting that 'I am here now' is not necessary in the usual sense—possibly, I'm not. From the other direction, assuming the Kaplanian framework, we can lay down semantic rules for expressions like N^i whose *content* will invariably be invariant, but where which invariant content is denoted is determined, in a context of use, by contingent matters at the context of use; unlike 'I am here now', whether or not such claims will be true depends on which context we utter them from.

As a result, there are two senses in which a claim like $N^i[\exists x x = x] \leftrightarrow \neg[\exists x x = x]$ ²² can be necessary: a sense corresponding to truth at all contexts of evaluation and a sense corresponding to truth at all contexts of use. These two senses are familiar from the distinction between deep and shallow necessity drawn by Evans (1979) and Davies and Humberstone (1980). Slightly simplified, a sentence ϕ is *shallowly necessary* if and only if:

$$\forall c \phi_{@}(c) = \mathbf{T}$$

That is, when the content of ϕ determined at the actual context @ is the constant function from contexts to true. A sentence ϕ is *deeply necessary* if and only if:

$$\forall c \phi_c(c) = \mathbf{T}$$

that is, when, for every context c , the content of ϕ determined at context c is true at c . Note that this distinction may, but need not, track any fancy view about

22. I have switched our leading example here to one constructed entirely out of logical expressions so as to connect this discussion cleanly with the results proved in the appendix.

actuality, content, and contexts as possible worlds. In order to make the relevant distinction, we merely need that there is a way to track contexts of use which fix the content of a notion and contexts of evaluation for that content.

Of course, given the way we have set up the criterion of logicality, there will be many true sentences which are composed only out of logical expressions, but which are not necessary. For instance, $\exists x, y x \neq y$, which expresses that there are at least two objects, is true but not shallowly necessary since there are at least two objects here, but there are contexts of evaluation containing only one object. Likewise, it is not deeply necessary since there are some contexts of use where it also fails. Yet it contains only content and character-invariant materials. This is a quite general problem, but one which we can finesse for now; our target will be to capture a sense in which claims composed only of logical materials are, if true, necessary. We will thus relativize the sense of necessity to contexts whose domains are isomorphic; i.e., we'll say that a sentence ϕ is *shallowly necessary_{inv}* if and only if:

$$\forall d[\phi_{@}(d) = \mathbf{T} \rightarrow \forall c \cong d \phi_{@}(c) = \mathbf{T}]$$

that is, when the content of ϕ determined at the actual context @ is true at a context d only if it is true at every context isomorphic to d . For deep necessity, we say a sentence ϕ is *deeply necessary_{inv}* if and only if:

$$\forall c[\phi_c(c) = \mathbf{T} \rightarrow \forall d \cong c \phi_d(d) = \mathbf{T}]$$

that is, when the content of ϕ determined at a context @ is true at @ only if the content of ϕ determined at any isomorphic context d is true at d .

Failure of shallow necessity_{inv} undermines logical truth (Hanson 2006). If a claim is not shallowly necessary, then its truth can vary from context to context, even when these contexts contain the same number of objects. And this kind of contingency seems to sit badly with the idea that logic is topic-neutral on one disambiguation of this claim—that is, that the truth of a logical truth is immune to swapping around facts about the relations with a context. Logical truths, especially those constructed from purely logical expressions, should be true in whatever context we evaluate them from, unlike 'there are orange cats'. However, certain statements which are not shallowly necessary_{inv} are still necessary in some sense, such as 'I am here now' or 'I exist'. These are true when they are evaluated in the context of use and are thus deeply necessary_{inv} in the sense defined above.²³

Plausibly, deep contingency_{inv} also undermines logical truth. If a claim is deeply contingent_{inv}, then its content varies from context of use to context of

23. This is to disagree with those who hold that logical truths can be shallowly contingent (Zalta 1988). The matter is too detailed for this paper, but see Hanson (2006) for what strike me as conclusive reasons to reject Zalta's claim.

use, even if its character stays constant. This kind of contingency sits rather badly with the idea that logic is topic-neutral on a different, but highly plausible disambiguation of this claim—that the content of a logical notion is immune to worldly facts *about* the context of use. Logical truths, especially those constructed from purely logical expressions, should be true *wherever* we assert them from, unlike ‘I am over six feet tall’ and potentially unlike ‘There are no unicorns’.

Call a claim that is necessary in both senses *strongly necessary_{inv}*. Truths constructed from character- and content-invariant materials are strongly necessary_{inv}. $N^i[\exists x x = x] \leftrightarrow \neg[\exists x x = x]$ is not. If its content is fixed in a K context, it is true and (shallowly) necessarily_{inv} so. From a $\neg K$ context, it is false and (shallowly) necessarily so. So it is shallowly necessary_{inv} and deeply contingent_{inv}. On the face of it, Hanson’s and McCarthy’s intuition that $N^i\phi \leftrightarrow \neg\phi$ is not intuitively necessary seems probative and corresponds to the failure of strong necessity_{inv}. True claims constructed from purely logical expressions should be *always* true—from any context of use, at any context we evaluate it at, modulo our above point about the size of the context—because their truth depends only on the meaning of logical expressions, no matter where the contents of these expressions are fixed.

We thus need an explication of topic neutrality which guarantees that the meaning (character) of the components of these truths are such that no matter what context we assert or define them, they turn out to be true at every point of evaluation *given that context*. We obtain this result by following the suggestion above, and requiring that logical expressions not only have invariant *contents*, but also invariant *characters*. (See the appendix for precise definition of these notions.) DOUBLE-STANDARD INVARIANCE—claimed here only as a necessary, not a sufficient condition—again does a far more plausible job at carving the logical from the non-logical.

3.1. Reconstructing Hanson and McCarthy’s Arguments

We can now reconstruct the modal argument against SIMPLE INVARIANCE: content-invariance, while entailing shallow necessity_{inv}, allows for deep contingency_{inv}, but logical truths should not be deeply contingent_{inv}. That is:

1. Invariance properties, as currently understood, apply to *contents* of expressions, not their meanings (understood as *characters*).
2. The character of an expression, understood as a semantic rule, can involve non-logical ingredients which still results in invariant contents.
3. Sentences involving expressions whose invariant contents are fixed by characters that involve non-logical ingredients can turn out to be deeply contingent_{inv}.

4. Logical truths should not be deeply contingent_{inv}.
5. Therefore, invariance of content is insufficient for logicality of an expression.

This argument avoids Sagi's objections to Hanson's and McCarthy's modal argument while staying true to the intuitive force of it.

Turning to the epistemic argument against the invariance criterion makes even clearer the role of character in accounting for logicality. *PseudoEnglish*₁ definitions of character like that of N^i are such that we can fully grasp them without being able to judge what content they specify. On the conditional interpretation of (N), we do not grasp the character of N —we merely grasp what the (constant) character would be given K and what the character would be given $\neg K$. But we can fully grasp the character of (N^i) without knowing what the content specified by that character is. If we do grasp what content is specified by that character, it is due to knowing some additional shallowly contingent materials such as K .

But grasp of the character of the logical constants—understanding their meaning—should allow us to suss out *a priori* that the logical truths generated from them are true, again presuming that we are in a position to know how many objects there are.²⁴ If we accept this as a constraint, then we obtain the following version of the epistemic argument:

1. Invariance properties, standardly understood, apply to *contents* of expressions, not their meanings (*characters*).
2. The character of an expression, understood as a semantic rule, can involve deeply contingent_{inv} materials.
3. If a sentence ϕ contains an essentially occurring expression γ whose character involves deeply contingent_{inv} materials, then even if the content designated is invariant, we may not be able to grasp that ϕ is true even if we fully grasp the character of γ and how many things there are.
4. The truth of truths constructed only from logical expressions should be derivable from a full grasp of the character of the essentially occurring constitutive expressions and a grasp of how many things there are.
5. Therefore, invariance of content is insufficient for logicality of an expression.

We should view these arguments as instructive in how to develop the invariance criterion in the less rarified context of non-extensional languages. If SIMPLE

24. This is essentially the complaint raised by Hanson (1997) against Sher's account. It also potentially explicates a sense in which logical truths are analytic.

INVARIANCE is insufficient for logicality in such contexts, as I have claimed along with Hanson and McCarthy, then we need something else to fill the gap. Invariance of character fills that gap nicely; hence, DOUBLE-STANDARD INVARIANCE looks to be the appropriate invariance-criterion for non-extensional languages.

4. An Alternative Approach: Sher's Rigidity Constraint

Gila Sher, in her seminal discussion of invariance criteria (1991), proposed that a conditions of being a logical constant is that it *rigidly* denote its extension:

Condition (B) ensures that logical terms are rigid. Each logical term has a pre-fixed meaning in the metalanguage. This meaning is unchangeable and is completely exhausted by its semantic definition. That is to say, from the point of view of Tarskian logic, there are no "possible worlds" of logical terms. Thus, qua logical terms, the expressions "the number of planets" and "9" are indistinguishable. If you want to express the intuition that the number of planets changes from one possible "world" to another, you have to construe it as an extralogical term. If, on the other hand, you choose to use it as a logical term (or in the definition of a logical term), only its extension counts, and this is the same as the extension of "9". (Sher 1991: 56)

Sher's stated aim is a characterization of logical constants in *logical* languages (2003: 197) and, in particular, logical languages of a Tarskian stripe. This motivates Sher to obtain rigidity by requiring that constants be defined by a single extensional function and identified with their extensions. Though Sher's condition works nicely for the languages she considers, the restriction to them seem less than perfectly motivated for a general account of logical constants. Why, for example, should we require that our metalanguage definitions of logical constants be extensional? Likewise, why only consider logical languages which can be so treated? Typical logical languages and accounts of logical expressions are often like this, but this fact should be explained. Anyways, our aim is the broader one of characterizing logicality for natural language. This may not be Sher's project, but is one for which her claim about rigidity is useful, if not perfectly apt.²⁵

25. Though it is not always entirely clear what her project is. In recent work, she claims that something loosely analogous to SIMPLE INVARIANCE gives a necessary and sufficient answer to the question "Which choice of logical constants will give rise to a logical system whose consequences transmit truth from premises to conclusion with an especially strong modal force in all fields of knowledge?" (2014: 182) But, if that is the question, then the restriction to extensional languages is very implausible since it is not at all obvious that we can formulate *adequate* logical systems for all fields of knowledge in extensional terms. DOUBLE-STANDARD INVARIANCE, absent this restriction, would seem to be a better answer to this question though perhaps not a sufficient condition.

Logical constants *should* be rigid in Sher's sense; non-rigid expressions, even when they have invariant extensions, are not intuitively logical.²⁶ Sher's point seems to be that expressions which change their meaning from world to world are intuitively non-logical expressions. Logical expressions should be rigid, but as we have seen above, once we move to the Kaplanian framework described above, there are two senses of meaning on offer. Rigidity captures the idea that the *content* of an expression is constant over worlds, but it does not capture the sense in which *character* is constant over context. 'I', famously, is a rigid designator of me; it denotes at every world, though in a different context, it would denote you in every world. Likewise, 'Nⁱ' is rigid, denoting either the trivial functor or negation, though which is a function of whether or not *K* is actually true. We seem to need here rigidity both of content *and* of character to capture the sense in which the meaning of logical expressions does not change.

So Sher's account, even if adequate for the particular logical languages she is considering, seems less adequate for our purposes than one which takes seriously her (and others, such as Feferman, McGee, and Hanson's) informal remarks about meaning and develops an independent criterion which underwrites her point about rigidity. Sher is aware of this limitation. She remarks in her (2001) that it matters little for her purposes, but significantly for other purposes, whether we ban *empirical* logical predicates—predicates which rigidly denote invariant extensions, but are defined by reference to contingent empirical materials.²⁷ However, adding a demand for character invariance to the account of logicity is better than banning such empirical predicates or treating them as a special case, even given her aims. It *explains* in a general and illuminating way why we should mark such predicates as non-logical.²⁸

As any character- and content-invariant expression will be rigid in Sher's

26. It is clear from her later discussion that Sher doesn't think that counterexamples involving constants which have different meanings on domains of different cardinality, such as McGee's (1996) funky disjunction, are non-rigid in her sense. Similar problems affect content- and character-invariant expressions, as discussed above (Section 3). As a referee points out, once we have identity and the existential quantifier as logical constants, finite cardinality quantifiers will be logical so these counterexamples don't have much force. Of course, this leaves non-finite cases, but anyways, it strikes me that so much as there is a problem here, it is one which should be solved by choosing a more appropriate class of transformation—such as those of Bonnay (2008) or Feferman (1999).

27. Her concessive remark is directed at the inability to know, *a priori*, the meaning of certain empirical logical predicates, not at the case of empirical materials playing a role in specifying the character of an expression.

28. Feferman has also argued, in two ways, for two-pronged approaches to logicity. In his (2010), he argues that logical notions should be *set-theoretically absolute* in order to avoid being beholden to a background account of set theory. In his later (2015), he argues that there should be a proof-theoretic analogue to the semantic notion of invariance. I am sympathetic with both claims, viewing them as arising from similar dissatisfaction with pure invariance criteria, but space precludes discussing them here. I hope to compare my approach with them elsewhere.

sense, but not conversely, the double-standard I am advocating seemingly improves on Sher's useful point above about rigidity in the context of natural languages. In fact, as the double-standard I advocate above implies rigidity in her sense,²⁹ we lose none of the cases Sher wants to exclude by means of her rigidity constraint, but additionally fend off intuitive counterexamples to Sher's view like N^i .³⁰ We can thus see DOUBLE-STANDARD INVARIANCE as a way of making sense of Sher's point in non-extensional contexts.³¹

5. Conclusion

Taking stock, I have argued that there are defensible versions of McCarthy and Hanson's modal and epistemic arguments against SIMPLE INVARIANCE, arguments that strongly suggest we ought to adopt DOUBLE-STANDARD INVARIANCE instead. The modal argument goes by way of claiming that invariance of content is insufficient to guarantee deep necessity_{inv}, but that deep necessity_{inv} should hold for any truth containing only logical expressions. The epistemic argument goes by way of showing that grasp of the character or semantic rule for logical constants is insufficient to determine *a priori* that truths constructed out of them are true. But this also seems required for logicity.

Deep necessity_{inv} and *a priori* discoverability from grasp of the character of an expression are, by themselves, insufficient for logicity. Consider 'I am here now'. This is deeply necessary—from any (normal) context, our assertion of this will be true. We can infer this fact *a priori* from our grasp of the meaning (character) of 'I', 'here', and 'now'. Likewise with other 'logical' truths of the logic of actuality and the logic of indexicals. But though 'I am here now' is deeply necessary, 'Necessarily, I am here now' is false. I might not have been. And logical truths should not be contingent in this way.³² Shallow necessity, as I have argued above, is arguably insufficient for logicity as well. A joint demand seems preferable since all truths constructed from standard logical expressions are necessary in both senses (subject to the '*inv*' modifier discussed above.)

29. Though, like expressions rigid in her sense, not sameness of denotation in every world unless we use a transformation other than isomorphism.

30. However, this is not to say that we can exclude all the cases which might be troublesome. See Footnote 34 below.

31. Sher also has a pragmatic reason to aim only at logical languages and to not care about empirical predicates as she is primarily interested in worldly *formal* structure. Space is too limited to discuss this here, but note that even if logical constants are formal in this sense, it doesn't follow that all formal notions can be grasped using only extensional resources. See, though, McGee (1996) for technical reasons to think that all standard logical objects can be defined using only extensional resources.

32. For a sensible discussion of logical truth in a Kaplanian context, see Michaelson (2014: 530–531).

For many mathematical and logical definitions, the distinction between character and content collapses. But once we look to define expressions for natural language, which contains more complex expressions such as indexicals, we can and should separate content and character. Once we have done so, it seems clear that simple content invariance is insufficient for logicity. And this I take to be at least part of the lesson of Hanson’s and McCarthy’s examples.

6. Appendix – A Sketch of Character Invariance

We sketch here the formal details of the notion of character invariance and the connection of invariance to necessity_{inv}, working with an unmodalized language for simplicity.³³ Let W be a non-empty set and D a function from W into a family of non-empty sets. We will call $w \in W$ a *world* (since it is our surrogate for ways things could have been) and D^w the *domain of w* . Let $D^* = \bigcup_{w \in W} D^w$. We build a type-hierarchy over D^* in the usual way. A *context* is a pair $\langle w, s \rangle$ of a member w of W and an object $s \in D^w$ (the speaker). Given a language L , we interpret expressions in L by assigning them contents according to the character of the expression, given a context. General extensions—*contents* in the sense we’ve been using—can be modeled by functions from W to members of the type-hierarchy built over D^* , appropriately to semantic type and member of W . The type of a unary predicate is, for example, a function from D^* to $\{\mathbf{T}, \mathbf{F}\}$, the type of a unary and monadic quantifier being a function from functions from D^* to $\{\mathbf{T}, \mathbf{F}\}$ to $\{\mathbf{T}, \mathbf{F}\}$, etc. A *character* is modeled by a function from contexts to contents.

So, for example, the character of ‘I’ is the function that takes a context $\langle w, s \rangle$ to the function from each member of W to s . The character of ‘cat’, being constant, is the function that takes a context $\langle w, s \rangle$ to the function from W to members of the type-hierarchy over D^* such that $w \mapsto f$ where:

$$f(a) = \mathbf{T} \Leftrightarrow a \text{ is a cat in } w.$$

The character of ‘=’, again being constant, is the function that takes a context $\langle w, s \rangle$ to the function from W to members of the type-hierarchy over D^* such that $w \mapsto f$ where:

$$f(\langle a, b \rangle) = \mathbf{T} \Leftrightarrow a = b.^{34}$$

Likewise, the character of ‘there are’, also being constant, is the function from W to members of the type-hierarchy over D^* such that $w \mapsto f$ where:

$$f(g) = \mathbf{T} \Leftrightarrow \{a \mid g(a) = \mathbf{T}\} \subseteq D^w \text{ and } \{a \mid g(a) = \mathbf{T}\} \neq \emptyset.$$

33. Various simplifications employed here make it simpler to give an idea of how to proceed, but the general suggestion is clearly portable to different semantical pictures.

34. The reader will note that this means that ‘ $a = b$ ’ can be true at w even if a does not exist at w . Such is life in modal semantics.

Finally, the character of N^i is not constant, being modeled by the function:

$$g(\langle w, s \rangle) = \begin{cases} v \mapsto \neg & K \text{ is true at } w \\ v \mapsto id_{\{\mathbf{T}, \mathbf{F}\}} & K \text{ is false at } w \end{cases}$$

As usual, when the function picked out is a characteristic function—i.e., one whose range is a subset of $\{\mathbf{T}, \mathbf{F}\}$ —we will treat it as the set; a unary predicate being the subset of D^* , a binary relation being a subset of $D^* \times D^*$, a unary quantifier being a subset of $\wp(D)$, etc.

Since all objects occurring in a domain of a world occur in D^* and we've drawn our contents from D^* , we will treat bijections between domains of worlds in terms of permutations on D^* (that there will be many permutations of D^* for each bijection from D^w to D^v will cause no problem given our purposes). Given a permutation π of D^* , we extend it to a function π^+ on all members of the type-hierarchy over D^* . We set $\pi^+(\mathbf{T}) = \mathbf{T}$, $\pi^+(\mathbf{F}) = \mathbf{F}$, and for $w \in W$, $\pi^+(w) = w$. For all d in D , let $\pi^+(d) = \pi(d)$. For an ordered n-tuple $\langle m_1, \dots, m_n \rangle$ of members of the type-hierarchy,

$$\pi^+(\langle m_1, \dots, m_n \rangle) = \langle \pi^+(m_1), \dots, \pi^+(m_n) \rangle.$$

Given a function f on types, $\pi^+(f)$ is the function composed of π^+ , f , and the inverse of π^+ : $\pi^+ \circ (f \circ \pi^{+^{-1}})$.

We say that a content co is isomorphism invariant on W and D when for any permutation π on D^* and pair of worlds w and w' such that $\pi^+(D^w) = D^{w'}$:

$$\pi^+(co_w) = co_{w'}$$

We say that co is simply isomorphism invariant when it is isomorphism invariant on W and D for any W and D . The content of the predicate 'stockbroker' is not isomorphism invariant. Given our actual world, $@$, the slightly better world w in which there are more revolutionaries and fewer stockbrokers and a permutation π on D^* such that $\pi^+(D^@) = D^w$, $\pi^+(\text{broker}_@) \neq \text{broker}_w$. On the other hand, it is easily verified that, for any permutation π , a pair $\langle s_1, s_2 \rangle \in \pi^+(= @)$ just in case $\langle \pi^+(s_1), \pi^+(s_2) \rangle \in =_w$, and so $=$ (the content of identity) is isomorphism invariant; this case, however, is borderline uninteresting as the content of $=$ is the same on every $w \in W$. A more interesting case is 'there is', and here it is again easily established that for any permutation π such that $\pi^+(D^w) = D^v$, $A \subseteq D^w$ just in case $\pi^+(A) \subseteq D^v$ and $A \subseteq D^v$ just in case $\pi^{+^{-1}}(A) \subseteq D^w$, so the set of non-empty subsets of D^v is the same as result of applying π^+ to the set of non-empty subsets of D^w .

We need an additional bit of apparatus for character invariance. Let an @-and-s preserving map from $\langle w, s \rangle$ to $\langle w', s' \rangle$ be a pair π, π_w such that

- (1) π is a permutation of D^* ,
- (2) π_ω is a permutation of W ,
- (3) $\pi(s) = s'$ and $\pi^+(D^w) = D^{w'}$,
- (4) $\pi_\omega(w) = w'$.

So π permutes D^* while preserving the speaker-role, while π_ω permutes the set of worlds, preserving the actual world role. We extend the pair π, π_ω to a single function π^+ in the obvious way.³⁵

We say that a character ch , as modeled in the set of contexts drawn from W and D , is isomorphism invariant for W and D when for any pair $\langle w, s \rangle, \langle w', s' \rangle$ ($w, w' \in W$ and s, s' in D^*) and @-and-s-preserving map π, π_ω :

$$\pi^+(ch_{\langle w, s \rangle}) = ch_{\langle w', s' \rangle}$$

That is, where applying π^+ to the image of ch under $\langle w, s \rangle$ — $ch_{\langle w, s \rangle}$ —results in $ch_{\langle w', s' \rangle}$. We say that a character for an expression ϕ is isomorphism invariant when no matter what W and D we choose, the resulting modeling of the semantic rule for ϕ as a function from contexts drawn from W to functions from W to the type-hierarchy over D^* is isomorphism invariant for W and D .

Now, given an @-and-s preserving map π, π_ω between $\langle @, me \rangle$ and $\langle @, Fred \rangle$, $\pi^+(me) = Fred$. $I_{\langle @, me \rangle}$ is the constant function from any $w \in W$ to me , so π^+ of $I_{\langle @, me \rangle}$ is the constant function from any $w \in W$ to $Fred$. But, of course, this is exactly $I_{\langle @, Fred \rangle}$. The invariance of the character of 'I' is borderline trivial given our definition. But this is as it should be—invariance of character tracks the fact that the semantic function of an invariant expression only depends on structural features of the context, such as the designated *speaker*. It does not matter who the speaker is, merely that there is one. We need @-and-s preserving maps here, not simply permutations of D^* , since setting character invariance up in terms of the latter would eradicate the possibility of invariance for notions like 'I' whose content differs depending on the choice of speaker.³⁶ The definition given above, of course, generalizes for larger indexes, though we will need a corresponding generalized notion of an index-preserving map. Here we will stick with the simple speaker-only index.

35. If we were to work with a modalized language, we would need to complicate this definition to accommodate additional modal expressions such as modal operators or the actuality operator @. As MacFarlane (2000) has pointed out, this requires taking a stand on how much modal structure—such as the accessibility relation—we should hold fixed. We ignore such complications here for expository purposes.

36. Note that the invariance of a term like 'I' depends on us allowing *any* member of the domain to count as a designated speaker or, alternatively, restricting the set of contexts to *proper* contexts in which the speaker of a context is, say, a person. The former is preferable; otherwise terms like 'I' carry significant non-logical information along with them—such as the information that the speaker is a person. Such terms are not intuitively logical. We will thus assume here that anything can serve as the designated “speaker” of a domain.

N^i has variant character. Let K be false at w and true at $@$ and π, π_ω an @-and-s preserving map between $\langle @, s \rangle$ and $\langle w, s \rangle$. $N^i_{\langle @, s \rangle}$ is the function such that $N^i_{\langle @, s \rangle}(v) = \neg$ for all $v \in W$. But $\pi^+(N^i_{\langle @, s \rangle})$ is $N^i_{\langle @, s \rangle}$ which is not identical to the function $N^i_{\langle w, s \rangle}$ —rather, $N^i_{\langle w, s \rangle}$ is the constant function from $w \in W$ to $id_{\{\mathbf{T}, \mathbf{F}\}}$. But both $N^i_{\langle @, d \rangle}$ and $N^i_{\langle w, d \rangle}$ are (trivially) content invariant. In contrast, ‘ Γ ’ has invariant character, but variant content. $I_{\langle @, me \rangle} = g$ where g is the constant function from W to me . Now, given π swapping only myself and Fred,

$$\pi^+(g_v) = \pi^+(me) = \text{Fred} \neq g_v$$

A little work shows that the usual logical constants, as well as standard cardinality quantifiers, have both invariant character and invariant content.

Call a statement *shallowly stable* just in case if it is actually true, it is shallowly necessarily_{inv} true, and if actually false, shallowly necessarily_{inv} false. Being constructed entirely from content-invariant expressions implies shallow stability_{inv}.

pf. Let ϕ be composed of only content invariant materials. Let w be a world and $\phi_w = \mathbf{T}$. Without loss of generality, we will presume ϕ is composed of two expressions, one of type S and one of type T , where S composes with T . Let v be a world such that $D^w \cong D^v$ and $\phi_w = S_w \circ T_w$. Since $D^w \cong D^v$, there is a permutation π of D^* such that $\pi^+(D^w) = D^v$. Let π be such. $\pi^+(S_w \circ T_w) = \pi^+(S_w) \circ \pi^+(T_w)$ (established by routine induction), so, since S and T are by assumption content invariant:

$$\pi^+(S_w \circ T_w) = \pi^+(S_w) \circ \pi^+(T_w) = S_v \circ T_v = \phi_v$$

Since $\phi_w = \mathbf{T}$, $\pi^+(S_w \circ T_w) = \mathbf{T}$, so $\phi_v = \mathbf{T}$. Similarly for \mathbf{F} .

Call an expression *deeply stable* just in case if it is actually true, it is deeply necessarily_{inv} true and if actually false, deeply necessarily_{inv} false. Being constructed entirely from character-invariant expressions implies deep stability:

pf. Let ϕ be composed only of character-invariant materials. Let $\langle w, s \rangle$ be a context such that $\phi_{\langle w, s \rangle}(w) = \mathbf{T}$. Without loss of generality, let ϕ be composed of two character-invariant expressions, S and T . Given a context $\langle w, s \rangle$, $\phi_{\langle w, s \rangle}(w) = S_{\langle w, s \rangle}(w) \circ T_{\langle w, s \rangle}(w)$. As before, for π, π_ω a @-and-s preserving map from $\langle w, s \rangle$, we have:

$$\pi^+[S_{\langle w, s \rangle} \circ T_{\langle w, s \rangle}](u) = \pi^+[S_{\langle w, s \rangle}](u) \circ \pi^+[T_{\langle w, s \rangle}](u)$$

If π, π_ω is a @-and-s-preserving map from $\langle w, s \rangle$ to $\langle v, t \rangle$ and ψ a character-invariant expression, then we have the special case:

$$\pi^+[\psi_{\langle w, s \rangle}](v) = \pi^+[\psi_{\langle w, s \rangle}(w)]$$

Since $\phi_{\langle w,s \rangle}(w) = \mathbf{T}$:

$$\mathbf{T} = \phi_{\langle w,s \rangle}(w) = \pi^+[\phi_{\langle w,s \rangle}(w)] = \pi^+[\phi_{\langle w,s \rangle}](v) = \pi^+[S_{\langle w,s \rangle}](v) \circ \pi^+[T_{\langle w,s \rangle}](v)$$

Since both S and T are character invariant:

$$\pi^+[S_{\langle w,s \rangle}](v) \circ \pi^+[T_{\langle w,s \rangle}](v) = S_{\langle v,t \rangle}(v) \circ T_{\langle v,t \rangle}(v) = \phi_{\langle v,t \rangle}(v)$$

So, stringing these identities together, we have

$$\phi_{\langle w,s \rangle}(w) = \phi_{\langle v,t \rangle}(v) = \mathbf{T}$$

and similarly for \mathbf{F} .

So if logical expressions are both character and content invariant—that is, if they are strongly invariant—then sentences composed entirely of them are both deeply and shallowly stable and hence strongly stable. Extending these methods to mixed sentences and to modalized language awaits another opportunity.³⁷

Acknowledgements

Thanks to Corine Besson, Denis Bonnay, Tim Button, Catharine Diehl, Paul Egré, William Hanson, Eliot Michaelson, Julien Murzi, Beau M. Mount, Gil Sagi, Florian Steinberger, and a couple of helpful referees for very valuable discussion. Thanks also to audiences at Leeds University, the School of Advanced Study at the Institute of Philosophy, University of Western Ontario, and the Munich Center for Mathematical Philosophy for very valuable feedback. The roots of this paper grew out of discussions with William Hanson during my time at the University of Minnesota. Thanks to him for stimulating my initial interest in this issue and being an early bulwark of sensibility against my more farfetched ideas.

References

Bonnay, Denis (2008). Logicality and Invariance. *Bulletin of Symbolic Logic*, 14(1), 29–68.

37. Note that notions like (N^W)

$$(N^W) \forall c [N^W \phi' \text{ holds at } c \leftrightarrow (\phi' \text{ holds at } c \text{ iff } \text{Water} \neq H_2O)]$$

where the substitute for K ($\text{Water} \neq H_2O$) is necessary *a posteriori* are also plausibly not character invariant. This allows us to explain away some troublesome like cases mooted by MacFarlane (2009). But it does not help with another troublesome set of cases, due to Gómez-Torrente (2003). Consider an analytically empty predicate like ‘male widow’. Such expressions plausibly are both content- and character-invariant, as defined, yet are intuitively not logical. It is some consolation that they are complex constructions, composed of character-variant materials, but the cases remain troublesome. I hope to address this in future work.

- Breckenridge, Wylie and Ofra Magidor (2012). Arbitrary Reference. *Philosophical Studies*, 158(3), 377–400.
- Button, Tim and Sean Walsh (in press). *Philosophy and Model Theory*. Oxford University Press.
- Davies, Martin and Lloyd Humberstone (1980). Two Notions of Necessity. *Philosophical Studies*, 38(1), 1–30.
- Dutilh Novaes, Catarina (2014). The Undergeneration of Permutation Invariance as a Criterion for Logicality. *Erkenntnis*, 79(1), 81–97.
- Evans, Gareth (1979). Reference and Contingency. *The Monist*, 62(2), 161–189.
- Feferman, Solomon (1999). Logic, Logics, and Logicism. *Notre Dame Journal of Formal Logic*, 40(1), 31–54.
- Feferman, Solomon (2010). Set-Theoretical Invariance Criteria for Logicality. *Notre Dame Journal of Formal Logic*, 51(1), 3–20.
- Feferman, Solomon (2015). Which Quantifiers Are Logical? A Combined Semantical and Inferential Criterion. In Alessandro Torza (Ed.), *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language* (19–30). Springer International Publishing.
- Gómez-Torrente, Mario (2003). The ‘Must’ and the ‘Heptahedron’. *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia*, 18(2), 199–206.
- Hanson, William H. (1997). The Concept of Logical Consequence. *The Philosophical Review*, 106(3), 365–409.
- Hanson, William H. (2006). Actuality, Necessity, and Logical Truth. *Philosophical Studies*, 130(3), 437–459.
- Jackson, Frank (1998). *From Metaphysics to Ethics*. Oxford University Press.
- Kaplan, David (1989). Demonstratives: an Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In Joseph Almog, John Perry, and Howard Wettstein (Eds.), *Themes from Kaplan* (481–563). Oxford University Press.
- MacFarlane, John (2000). *What Does It Mean to Say That Logic Is Formal?* (Unpublished doctoral dissertation). University of Pittsburgh.
- MacFarlane, John (2002). Frege, Kant, and the Logic in Logicism. *Philosophical Review*, 111(1), 25–65.
- Mautner, Felix (1946). An Extension of Klein’s Erlanger Program: Logic as Invariant-Theory. *American Journal of Mathematics*, 68(3), 345–384.
- McCarthy, Timothy (1981). The Idea of a Logical Constant. *The Journal of Philosophy*, 78(9), 499–523.
- McGee, Vann (1996). Logical Operations. *Journal of Philosophical Logic*, 25(6), 567–580.
- Michaelson, Eliot (2014). Shifty Characters. *Philosophical Studies*, 167(3), 519–540.
- Perry, John (1979). The Problem of the Essential Indexical. *Noûs*, 13(1), 3–21.
- Sagi, Gil (2015). The Modal and Epistemic Arguments Against the Invariance

- Criterion for Logical Terms. *Journal of Philosophy*, 112(3), 159–167.
- Sher, Gila (1991). *The Bounds of Logic: A Generalized Viewpoint*. Cambridge University Press.
- Sher, Gila (2001). The Formal-Structural View of Logical Consequence. *Philosophical Review*, 110(2), 241–261.
- Sher, Gila (2003). A Characterization of Logical Constants is Possible. *Theoria*, 18(2), 189–198.
- Sher, Gila (2013). The Foundational Problem of Logic. *Bulletin of Symbolic Logic*, 19(2), 145–198.
- Stang, Nick (2014). Kant, Bolzano, and the Formality of Logic. In Sandra Lapointe and Clinton Tolley (Eds.), *The New Anti-Kant: Translation and Essays* (192–234). Palgrave.
- Tarski, Alfred (1936). Über den Begriff der logischen Folgerung. *Actes du Congrès International de Philosophie Scientifique, VII Logique, Actualités scientifiques et industrielles 394, Paris* (1–11).
- Tarski, Alfred (1986). What Are Logical Notions? *History and Philosophy of Logic*, 7(2), 143–154.
- Tolley, Clinton (2013). Kant on the Generality of Logic. In Margit Ruffing, Claudio La Rocca, Alfredo Ferrarin, and Stefano Bacin (Eds.), *Proceedings of the 11th International Kant Congress* (Vol. 2, 431–442). De Gruyter.
- Woods, Jack (2014). Logical Indefinites. *Logique et Analyse*, 57(227), 277–307.
- Zalta, Edward (1988). Logical and Analytic Truths that Are Not Necessary. *The Journal of Philosophy*, 85(2), 57–74.