ON WORDS*

Words strain,
Crack and sometimes break, under the burden,
Under the tension, slip, slide, perish,
Decay with imprecision, will not stay in place,
Will not stay still.
—T. S. Eliot, *Quartet I, Burnt Norton*, 1935

In his seminal paper "Words,"[1] David Kaplan addresses a pair of questions that have been largely neglected by the philosophical community:

(i) Under what conditions are two utterances utterances of the same word?
(ii) What *are* words?

That these questions have not received much attention is rather surprising: after all, philosophers and linguists frequently appeal to considerations about word and sentence identity in connection with a variety of puzzles and problems that are foundational to the very subject matter of philosophy of language and linguistics.[2] Kaplan's attention to words is thus to be applauded. And there is no doubt that his discussion contains many useful insights. Nevertheless, we find his picture deeply flawed for a variety of cross-cutting reasons. Our aim in this paper is to further advance an understanding of the nature of words, both by remedying the problems with Kaplan's account,

* Earlier drafts of this paper were presented at the Rutgers Semantics Workshop; the World Congress in Philosophy in Seoul, Korea; Saint Petersburg State University; University of California, Santa Cruz; University of Oslo; University of Paris; and the American Philosophical Association Pacific Meetings in April 2010. We would like to thank the audiences of each for their input, especially David Kaplan and Sylvain Bromberger. We also thank the anonymous referees of this JOURNAL for their helpful comments.

We are most grateful to David Kaplan for his extensive, insightful, and witty commentary at the 2010 Pacific APA on sections I–III of our paper. We have made a few adjustments where we were convinced that we had not paid sufficient attention to his original text. We have flagged a few other places where his public remarks gave us additional evidence concerning his original intentions. We have tried for the most part to make the adjustments in footnotes; in particular, we added or altered footnotes 12, 20, 22, 26, and 40.

[1] Kaplan, "Words," *Proceedings of the Aristotelian Society, Supplementary Volumes*, LXIV (1990): 93–119. Subsequent references to this work appear in the text.

[2] See, for example, the discussions of Motivations Three through Five in section III below. The importance of word individuation to mainstream philosophy is not new; notice, for example, its relevance to those varieties of nominalism that analyze the phenomenon of having something in common in terms of *the same word* being truly predicable of many objects.

0022-362X/11/0000/001–039 © 2011 The Journal of Philosophy, Inc.

and also by achieving a suitable perspective on what the metaphysical investigation of word identity can hope to achieve.

Our discussion divides into four sections. In section I, we evaluate Kaplan's discussion of a contrast integral to his own account: that between the type-token and the stage-continuant conceptions of words. In section II, we present three constraints on any account of words and two further themes in Kaplan's discussion central to his conception of words—the role of repetition and the constitutive authority of intentions. While these ideas have laudable motivations, we argue they are far from the best way of making good on the insights that drive them. The final two sections take a skeptical turn. In section III, we express doubt about Kaplan's presumption of the importance of what he calls 'common currency names', thus raising a suspicion that he may be in pursuit of a chimera. Finally, in section IV, we express pessimism about whether interesting answers to question (i) above will be forthcoming. Does the legitimacy of our word ontology need or depend on the availability of such answers? We explore these issues in some detail, mindful of the fact that the case of words is in many ways representative of many other ontological categories from the manifest image. Along the way, we tease apart a number of metaphysical questions in the vicinity of the topic of word individuation—questions rarely disentangled—and consider how the discussion of the previous parts bears on them.

I. TYPE-TOKEN VERSUS STAGE-CONTINUANT

There are physical events in spoken language—utterances. There are physical products of written language—inscriptions. There are physical events in sign language—signings. And there are words that certain utterances, inscriptions, and signings are utterances, inscriptions, and signings of. What exactly is the relation between the former class of events and objects and the words themselves? Kaplan presents us with two candidate models.

On the first model of words, the *type-token model*, a word is an abstract object instantiated by various physical objects and events—*tokens*. Further, as Kaplan conceives it, the type-token model treats the abstract object as encoding a certain form that is common to the tokens. Tokens thus get to be tokens of a word by having a physical embodiment with the right intrinsic profile. On the simplest and most common view, word identity begins with orthographic and/or phonetic shape,[3] that

[3] Donald Davidson, "Quotation," in *Inquiries into Truth and Interpretation* (New York: Oxford, 1979), pp. 79–92, at p. 90. Compare with Susan Haack, *Philosophy of Logic* (New York: Cambridge, 1978), p. 75: "[Expressions are] either a pattern which similar

is, with spelling and pronunciation. Two words are identical just in case they share the same relevant shape.[4]

On the orthographic side, Wetzel tells us that "'dog' is a different expression from 'god'." Not only the letters, but the *order* of the letters occurring in an expression determines the expression's identity."[5] Orthographically, then, a word is identified with a sequence of letters with a space at each end and no space in its middle.[6]

Linguists, unlike philosophers and lexicographers, ignore orthography altogether and concentrate on sound, in particular, on sequences of phonemes. When Fiengo and May say, "…phonological distinctness, when it is believed to fall outside the range of dialectal variation, may be individuative of words,"[7] they intimate a partial discrimination of words through sound, as does Bromberger when he writes of the *word* 'dog' that it "is monosyllabic."[8] Segal and Speas reject Davidson's claim that the 'that' of indirect quotation is the same word as the orthographically indistinguishable demonstrative word 'that' by noting that their phonologies are distinct.[9]

Concentration on pronunciation as the identifying feature of words extends back to Aristotle, who treated written words as mere 'stand-ins' (or 'subordinates' (Ockham)) for spoken ones. However, some have noted (Pierre d'Ally) that a written message can be understood directly, without recourse to sound; there are so many words we know how to spell but not pronounce.

---

tokens exemplify, or a class of similar tokens"; and Philip Hugly and Charles Sayward, "Expressions and Tokens," *Analysis*, XLI, 4 (October 1981): 181–87, at p. 184: "expressions…can be construed as classes of perceptible particulars similar in some physical respect to given perceptible particulars."

[4] "An expression is a phonic or graphic form: the former has acoustic properties and the later spatial ones." See Brendan S. Gillon, "Ambiguity, Indeterminacy, Deixis, and Vagueness: Evidence and Theory," in Steven Davis and Gillon, eds., *Semantics: A Reader* (New York: Oxford, 2004), pp. 157–87, at p. 162. Compare with John Lyons, *Semantics: Volume 2* (New York: Cambridge, 1977), p. 18.

[5] Linda Wetzel, "What Are Occurrences of Expressions?" *Journal of Philosophical Logic*, XXII, 2 (April 1993): 215–19, at p. 218.

[6] W.V.O. Quine, *Word and Object* (Cambridge: MIT, 1960) recognized that "the principle behind the printer's use of spaces is dim," p. 13. What do we make of expressions like 'ice cream' and 'ice-cream' or, for that matter, languages with writing systems without spaces—as with most ancient languages, including Phoenician, Greek, and Latin, and Medieval Latin? Yet—as do most philosophers—Quine acquiesces to the orthographic account (*ibid.*, p. 14).

[7] Robert Fiengo and Robert May, *Di Lingua Belief* (Cambridge: MIT, 2006), p. 54.

[8] Sylvain Bromberger, "Types and Tokens in Linguistics," in Alexander George, ed., *Reflections on Chomsky* (Malden, MA: Blackwell, 1989), p. 85.

[9] Gabriel Segal and Margaret Speas, "On Saying That," *Mind and Language*, I, 2 (June 1986): 124–32.

According to either orthodoxy, then, nothing in a word weds it to any particular grammatical category, meaning, or even a specific language. The same word (that is, sounds or shape) might be both a noun and a verb, carry several or even no meanings, co-exist in different languages, or change any of these features across time and location.[10]

On the second model of words, the *stage-continuant model*, which Kaplan prefers, "utterance and inscriptions are *stages* of words, which are the *continuants* made up of these interpersonal stages" (98). The phrase 'made up' implies a compositional relation. The picture conveyed is that words are long-lived objects composed out of short-lived objects—utterances and inscriptions. Given these remarks, it is natural to elucidate the stage-continuant model using a framework that is standard to four-dimensionalist metaphysics: a word is a four-dimensional continuant with various utterances or inscriptions as its short-lived stages.[11] A particular utterance or inscription of a word *w* thus turns out to be a spatiotemporal part of *w*: *w* is a fusion of the archipelago of its spatiotemporal parts. Of course, as Kaplan realizes, this still leaves many questions unanswered.[12]

The stage-continuant conception does not, in and of itself, tell us when a fusion of utterances or inscriptions counts as a word, and it does not tell us when two utterances or inscriptions are parts of the same word continuant. But this is as it should be. By analogy, the stage-continuant model of personal identity, according to which people are fusions of person stages, does not tell us under which conditions two stages belong to a single person continuant and under which conditions a fusion of person stages counts as a single person. But both at least purport to provide a framework within which the metaphysics of persons and words can be suitably pursued.

First, we wish to emphasize that we quite agree with Kaplan that a philosophically satisfying theory of words cannot proceed entirely within a shape- or form-theoretic framework. Simple reflection on the fact that the same word can be written, uttered, signed,[13] Brailled,

[10] Compare with Simon Blackburn, *Spreading the Word* (New York: Oxford, 1984), p. 20.

[11] Truths to the effect that a certain word undergoes changes will be accommodated by such a conception by standard four-dimensionalist accounts of what change over time comes to.

[12] Kaplan's APA presentation indicated that while, to our mind, his written exposition suggests the four-dimensional conception, he did not really intend to be endorsing a metaphysical picture of that sort. His presentation also made clear that he intended the stages of words to include the internal token representations by means of which a word is stored (even though his actual gloss on the nature of word stages does not make this intention manifest).

[13] We have in mind Signed English, a system of manual communication that strives to be an articulation of English vocabulary and grammar, in contrast to American Sign Language, a separate language with its own vocabulary and grammar.

or semaphored already renders such a conception dubious: after all, there is hardly anything shape- or form-like in common among an utterance, an inscription, a hand gesture, and a bump on a panel. In sum, the form-theoretic conception tends to ignore the fact that a variety of systems can articulate a single word. As Kaplan emphasizes (see, in particular, 102–06), even if we restrict attention to a particular system of articulation, the form-driven conception does not square with the fact that the way a word is articulated within that system can change *its* shape or form significantly over time and location. The word 'color' has two "modern current or most usual spellings" ['colour', 'color'] and eighteen historical ones.[14] Accordingly, 'color' and 'colour' are two spellings of the same word.

We do not deny that there may be *certain* contexts in which we use 'word' to speak about entities that are individuated by shape or form. After all, there are contexts in which we can communicate truths with such things as "The word 'color' has exactly five letters," "The words in that book are a funny shade of gray," and so on. This kind of flexibility is not particularly surprising. Consider the flexibility of our uses of 'that book': sometimes we use it to indicate a particular token volume; sometimes a particular work, say, Dante's *Inferno*; sometimes a particular edition ("I wish I owned Aldus's edition of Dante. I have two friends that own that book.").

The form- or shape-theoretic conception *may* capture some of our uses of 'word' (we shall return to this topic below). But there *are* clearly uses for which nothing like that conception is adequate. Like Kaplan, we are interested in these latter uses.

(The form-theoretic conception is no doubt driven by the insight that a word's meaning is inessential to it. Crudely put, the key idea is that if the meaning is inessential, then what is left except form? As Kaplan's own discussion renders clear, though, this thought turns on a failure to realize that relational factors may be crucial as to whether an utterance or inscription is an utterance or inscription of a word. We return to this issue in due course.[15])

[14] *Oxford English Dictionary*, 1971 ed., vol. 2, s. v. "color."

[15] In a classic study, Peter Ladefoged and D. E. Broadbent manipulated the formant frequencies of material preceding a target word that could be perceived as 'bit', 'bet', 'bat', or 'but'. These manipulations to the material affected the way listeners categorized the final target word, producing context effects such that the target word was identified differently as the formant frequencies of the preceding material changed. The spectral manipulations to formant frequencies of the precursor material created a perceived change in voice characteristics, making it sound like different speakers uttered the target. See Ladefoged and Broadbent, "Information Conveyed by Vowels," *The Journal of Acoustical Society of America*, XXIX, 1 (January 1957): 98–104.

The second—more critical—point we wish to register is that Kaplan presents us with a false dichotomy. (We concur with McCollough, though we disagree with him that all tokens of a word must be "similar or resemble" each other in an interesting intrinsic respect.[16]) One can favor the type-token conception in thinking of a word as an abstract object without taking on board any commitment to a shape- or form-theoretic conception of words. Consider, by analogy, species: in repudiating a view of species as four-dimensional objects, one need hardly be committed to a shape-or form-theoretic view of what species are. The barebones abstract-object view of species merely tells us that a species is a nonconcrete object and that there is a relation which holds between a particular creature and its species—the *of*-relation. Nothing requires that the conditions for the of-relation must hold in virtue of the intrinsic features of the creatures. In particular, it is quite compatible with the barebones view that historical-relational features are crucial to whether an individual belongs to a particular species.[17]

We conclude that even when the words-as-forms view is repudiated, a real choice of metaphysical framework remains: we might, in line with the picture suggested by Kaplan's stage-continuant gloss, think of words as four-dimensional objects which have inscriptions or utterances as temporal parts on the stage-continuant model.[18] Or, we

---

We interpret this experiment as evidence against the view that the intrinsic character of an utterance determines whether it is an utterance of a certain word. As we mention in the text, there may be uses of 'word' which are attuned to form. These reflections show that if such conceptions are not bankrupt, founded on illusions of commonalities that are not present, then even they cannot be explicated by appeal to intrinsic character alone. It is also not hard to make trouble for an orthographic conception that presumes that, in the relevant sense, orthographic form is intrinsically determined. Note, for example, that a physical space appears between the articulated letters of a single word even though these do not realize a space in the orthographic sense. Our judgment as to whether a physical space constitutes a space in the orthographic sense depends upon the relative size of that space to other spaces. If the gap between the letters 't' and 'o' is half an inch but flanked by spaces of two inches, we are not deterred from treating the sequence 't' and 'o' as expressing a single word. This all makes further trouble for those who think that the intrinsic character of an expression determines whether or not it articulates a word, since the relative-size facts just adverted to are not intrinsic to the expression itself.

[16] Gregory McCulloch, "Making Sense of Words," *Analysis*, LI, 2 (March 1991): 73–79, at p. 74.

[17] Kaplan, *op. cit.*, p. 98 seems to think the stage-continuant model is required in order to accommodate historical-relational insights. But he gives no justification for this contention.

[18] We note that one might wish to allow that one temporal part of an inscription is a temporal part of one word and a later temporal part is not part of that word—just as an earlier but not later temporal part of an atom might be part of a person.

might just as well treat a word as a nonconcrete object which is articulated by various concrete events or objects. We shall call this the *abstracta-articulations model*.[19,20]

We have two general reasons for preferring the abstracta-articulations model over the stage-continuant model. Our first concern with the stage-continuant model turns on a more general choice point of modal metaphysics. It seems clear enough that various facts about the performance profile of a word are *in*essential to the word. Uses of a given word could have become obsolete. It once might have been articulated more loudly; and so on. Supposing a word to be an archipelago of utterances and inscriptions, how are these sundry facts to be accommodated? One particular source of disquiet is that insofar as we think of the continuant as an archipelago of utterances, our modal intuitions may not at all match the modal profile with which the word was originally associated. This concern is familiar from discussions of four-dimensional objects in general: supposing that a person is a fusion of person stages, and given that there are ways of thinking about aggregates of stages which yield very different modal intuitions, how do we accommodate such facts as that a person could have lived for a much shorter or longer time?[21] The standard mode

[19] We ourselves would have no deep objection to using 'type-token' language for this model, but since Kaplan insists on a shape-theoretic construal of the language of types and tokens we shall stay clear of it.

[20] There are yet other models that might be considered. As Mark Johnston emphasized to us (personal communication), one might opt for the view that words (and species) are higher-order persisting objects that are concrete; that are to be sharply distinguished from their realizations in utterances, inscriptions, and so on; and that—*contra* the stage-continuant model—are best thought of within the framework of three-dimensionalism. On this view, as against the abstracta-articulations model, words only exist at a world insofar as some first-order realization of them does. At no time can a word exist without some physical object constituting it. Meanwhile, as against the stage-continuant model, words should not, on this view, be thought of as composed of the multitude of their tokens. Rather, at any particular time there will be certain tokens that constitute it at that time. As will be clear from our discussion, our own current preference is for the abstracta-articulations model. But we recognize that the space of positions ought not to be limited to those described in the main body of the text. In general, those who wish to think of words as concrete will have more intellectual wiggle room insofar as they allow that two concrete objects can occupy the same four-dimensional region of space-time. We lack the space to pursue these issues further here. (We note that in discussion Kaplan expressed a preference for something like Mark Johnston's view, as opposed to the stage-continuant view on its four-dimensionalist construal. We also acknowledge the need for further clarification of what the purported three-dimensionalism/four-dimensionalism contrast comes to.) For more, see John Hawthorne, "Three-Dimensionalism," in *Metaphysical Essays* (New York: Oxford, 2006), pp. 85–110.

[21] A standard puzzle in the vicinity is the following. Take the worm that is a certain person x and the worm that is the temporal part p of x that spans the first thirty years of x's life. Suppose x lives sixty years. It is possible that x could have lived only thirty

of accommodation in person-as-continuant metaphysics is via counterpart theory: various possible fusions stand in the person-counterpart relation to the actual continuant, and various ascriptions of essential and accidental properties to the continuant are rendered true or false by the profiles of the possible fusions that are person counterparts of that continuant.[22]

We anticipate that insofar as one pursues a continuant model of words, one will likely be drawn into an analogous account of attributing modal ascriptions to words. Crucially, however, the viability of the approach will turn on the viability of the counterpart-theoretic approach to *de re* modal judgments. While this is not the place to pursue these more general issues, our inclination is to be wary of counterpart theory and, for that reason, to be wary of the stage-continuant model. Insofar as a reader shares our suspicion of counterpart theory, he should at least be sensitive to issues as to how the stage-continuant model is supposed to accommodate modal platitudes without immersing itself in an objectionable modal meta-semantics.[23]

Our second concern with the stage-continuant model of words is more straightforward and ought to have purchase whether or not one is a fan of counterpart theory. It is clear enough that Kaplan is committed to the view that all words are articulated in one way or another. ("They live in the world, not in Plato's Heaven" (111).) If a word is an archipelago of articulations, then absent a single articulation there is *no* word. But this is an untoward result. Consider, for example, the fact that the prefix 'un-' ('im-', 'in-', 'il-') can combine with ever so many adjectives to produce a word.[24] It seems evident

---

years and had a life that matched x's actual life. It is natural to think that at such possible worlds p exists. By hypothesis, at such worlds x exists. But assuming the classical extensional mereology standard among four-dimensionalists, we are then under pressure to say that x and p are identical at those worlds, violating the necessity of distinctness. For relevant discussion, see Peter Van Inwagen, "Four-Dimensional Objects," *Noûs*, XXIV, 2 (April 1990): 245–55. A second puzzle is this: when we think of a person as a concertina of stages, it is *not* natural to think that it could have been very short in temporal duration, while when we think of the person as a person, it is natural. Which perspective is to be privileged? For relevant discussion, see David Lewis, "Counterparts of Persons and Their Bodies," this JOURNAL, XLVIII, 7 (April 1971): 203–11. The same puzzles arise in the current setting, *mutatis mutandis*. Counterpart theory offers perhaps the most promising way of navigating such puzzles to advocates of temporal parts who are wedded to standard mereology.

[22] Obviously, counterpart theory is not Kaplan's preferred way of thinking about modality. The point is that the picture suggested by the thesis that words are 'made up' out of their stages sits best with counterpart theory.

[23] See Michael Fara and Timothy Williamson, "Counterparts and Actuality," *Mind*, CXIV, 453 (January 2005): 1–30.

[24] Gillon, *op. cit.*, p. 164.

that a particular combination of that prefix and an adjective—as in 'unhappy', for example—might never have been articulated.[25] But that would hardly preclude this combination from being a word of the language. It may even be that in this scenario many people thought about using the word 'unhappy' but decided against it, or intended to use it but never got around to it.[26] Given this presumed profile of thoughts and intentions, it would be very awkward to deny that the word exists. And yet it is very hard to see how it could exist once a Kaplanian framework is adopted. Indeed, the case could be made stronger; once one realizes that some morphology is productive, as in 'anti-missile', 'anti-anti-missile', and so on, it becomes clear that there are more unspoken words than spoken ones (*contra* Kaplan, who insists, "The world is not brimming with unspoken words" (117)).

The Kaplanian might try to retreat by claiming that the stage-continuant model is only appropriate for words that are not built up out of other words, and propose a different account of the remainder.[27] But this is inadvisable. For one thing, we surely would prefer a more unified account if we could get it. Second, such a retreat fails to appreciate the import of the word-morpheme distinction. 'Un-' is not a word; it is a morpheme.[28] In the imagined scenario, the unarticulated existence of 'unhappy' as a word cannot be explained via

[25] One might object, as does Jeff King (personal communication), that since morphology is not fully systematic—it is not in general true that any combination of 'un-' with an adjective yields a word—it makes no sense to suppose that 'unhappy' is a word and yet lacks concrete realization. Note that there are languages where morphology is much more systematic (Mohawk is one; thanks to Mark Baker), and hence, where this objection would not have purchase. This worry has less force against certain other examples from English—see our 'anti-anti-missile' example below. In that case, the application of a rule for 'anti-' is clear even in the absence of the relevant concrete instance. Note also—deploying a theme from the main text—that one could stipulate that 'un-' does combine with 'happy' without actually concretely realizing 'unhappy'. In that case, the anticipated complaint has much less force.

[26] Of course, if internal storage counts as articulation (cf. note 13), then a time when people thought about using 'unhappy' may, on a natural construal, count as a time when it was articulated, albeit internally. While this appeal may help the stage-continuant view in some cases, it will not provide the resources to block the unwelcome conclusion that, at times when the particular combination of morphemes is not even internally articulated, the relevant word passes out of existence. Moreover, if one thinks that when a word stops being articulated it stops existing and *is incapable of coming back into existence*—in conversation, it became clear that this is an important part of Kaplan's anti-Platonism—then one risks having to say that a new word 'unhappy' gets considered whenever there is a lapse during which the relevant combination of morphemes never gets articulated.

[27] As far as we can see, there is no anticipation of such a retreat in the original paper.

[28] Words are distinct from morphemes. They are the smallest units of a sentence with positional mobility, meaning roughly, they are the smallest units of a sentence redistributable *salva congruitate*. The words 'men' and 'women', for example, display their distributional freedom in (a)–(b).

the existence of two simple words, 'un-' and 'happy', since this would misclassify 'un-'. (The same can be said for the morpheme 'anti' in the word 'anti-missile'.)

It is *also* perfectly possible that something *un*articulated counts as a word because it is an acceptable combination of morphemes, none of which is itself a word. But the anticipated retreat lacks the resources to accommodate this scenario. To reinforce this point, one need only consider possible facts about word invention. Suppose the impositor proclaims, "The sequence of morphemes 'un-' followed by 'voke' is to mean such and such." It is intuitively clear that his speech serves to *introduce* the word 'unvoke' into the language. Suppose soon after his decision, the world blows up. The word 'unvoke' has been introduced by a description and so has become a constituent of the language, but it has never been *used* in discourse. Notice, in particular, it did *not* occur in the speech that served to introduce it.[29]

(a) Do men like women?
(b) Women men like.

Morphemes, on the other hand, have a rather fixed sequential order, as evidenced in (c)–(d).

(c) unhappy
(d) *happyun

Syntax can manipulate words but not smaller units. As a corollary, it is true that there are 'free word-order' languages (to varying degrees, including English), but no 'free morpheme-order' languages.

Words, unlike morphemes (and complex expressions), resist linguistic material insertion. (a) permits the insertion of an indefinite number of new items, as in (e)–(f):

(e) Do men who are single like women who are single?
(f) Do men often like women?

More generally, words can be inserted between words, and morphemes can be inserted between morphemes (when conditions are right), but words cannot be inserted between morphemes. For example, we say of (g) that 'the' is a word and '-s' is a morpheme partly because we can insert words between 'the' and 'boy' but not between 'boy' and '-s', as in (h) and (i):

(g) the boys
(h) the tall boys
(i) *the boy responsible s (though 'the boys responsible' is fine)

Another large issue here is phonological: words are self-contained phonological units; morphemes are not. Words have at least one syllable, their own stress, and so on. Morphemes need not (see again plural '-s'). (Thanks to Mark Baker.)

[29] Again, matters become more delicate if we allow articulations to include internal articulations. For in the situation described, it is natural to think that the word 'unvoke' is internally represented. In his APA presentation, Kaplan wished to (a) require that there be such an internal articulation in order for the word to come into existence and (b) insist that once all articulation lapses then the word permanently goes out of existence. We think that there is a strong case against (b). Suppose we introduce

The abstracta-articulations model faces no problems in accommodating word-introduction scenarios of this sort. But they embarrass the stage-continuant model, since there are no stages upon which to ground the existence of the word. (We similarly can introduce a written word descriptively by saying which sequence of letters comprises it. If we say, "Let the sequence of letters 'L' followed by 'U' followed by 'V' mean *love*," though we have introduced a word, we have not yet tokened it.)

Of course, the abstracta-articulation model for words is logically compatible with a stage-continuant model for morphemes. Might the Kaplanian theory be salvaged by reconfiguring it as a view about morphemes? Such a view handles some of the problems registered in the text, but not the last one raised in the above paragraph.[30]

(Consideration of complex words also suggests the following, admittedly more recherché, worry for the stage-continuant model. There could be a language where one constructs a word w_2 by placing a word w_1 after itself, and where the only uses of w_1 are as parts of an articulation of w_2. For example, it might have been that 'people' was only ever used in constructions like "They are people people." Perhaps all sorts of other uses are allowed but as a brute contingent fact that potentiality is never actualized. In this scenario, a fusion of articulations of w_2 would appear to have exactly the same four-dimensional profile as a fusion of articulations of w_1. If one embraces

---

'unvoke', use it for a while, and then forget about it. No articulation remains. A few years later we come across a video that depicts the original introduction speech. It captures the moment where the speech "Let the sequence 'un' followed by 'voke' be a word that means…" is made. Having witnessed the video we pick up using the word again. (b) will reckon such a description of the case incoherent: one will be forced to think of a new word coming into being when the video is witnessed. Notice it is no defense to say that since the video was in existence all along there was an articulation of the word all along. This is to strain the notion of articulation beyond any useful sense. Given that the original overt description did not articulate the word, it is clear that in no reasonable sense does the video of that speech do so. (We note in passing that it is far from clear in any case that the picture that words are concrete changeable objects ought itself to deliver (b). One might think that the very same boat or watch may exist at two times in between which there is a time when it does not concretely exist owing to its being disassembled for winter storage or for repair by the watchmaker. And surely no one needs words to be more concrete than boats and watches.) One test case for (a) is where the structure of internal articulation matches that of the overt speech: there is an internal description but in no reasonable sense an internal realization of the word. We shall not explore this or related test cases further here.

[30] Note that if Kaplan's view were reconfigured as a view about morphemes, it might put pressure on what he says about the constitutive role of intentions, since ordinary intentions to repeat do not seem to have morphemes as their objects.

a unique fusion principle for continuant mereology, one will be forced to admit the untoward result that w_1 and w_2 are the same word!)[31,32]

II. THREE CONSTRAINTS AND TWO PROBLEMATIC IDEAS

Much of what Kaplan says about the stage-continuant model carries over to the abstracta-articulations model. Hence, the discussion in section I leaves many evaluative questions about Kaplan's discussion unresolved. In this section, we first identify a number of themes that shape his discussion and which we shall also adopt as framing

[31] The unique fusion principle says that at most one thing fuses a given class of objects. The intuitive idea is that one disallows two objects made of the same bits.

[32] Given that morphemes are abstract objects and that the two putative occurrences seem to involve a numerically identical morpheme, what does it mean, exactly, to say one morpheme occurs *twice* in a word ('anti-anti-pacifist', for example)? We will make a few observations in connection with this constituency question:

(i) Do *not* think of words as built up out of letters. That conception of words privileges written articulations of words in an unjustified way, since there is no good reason to think of spoken words as built up out of letters.

(ii) The problem should be treated on all fours with a variety of intimately related problems: what does it mean to say that the proposition *that two is even* 'occurs twice' in the disjunctive proposition *that two is even or two is even*? What does it mean to say that there are two hydrogen molecules in H_2O? In each case it is not clear how to resolve the issue with a mundane appeal to type-token language, since the relevant constituents seem to be abstract objects rather than 'tokens' of them. Issues of this sort have frequently been discussed by metaphysicians under the heading 'structural universals', and we cannot do full justice to the range of issues involved here. But we can gesture at our own preferred solution.

Often, what postures as counting of a certain sort of object is, at bottom, a mode of counting a number of states and events that involve that object. Thus, recalling a standard example, we may say that twelve barges came through a lock on a particular canal on a particular day, even where this involves three physical barges passing through four times each. What 'twelve' is a count of is not some physical barges, but instead, plausibly, of twelve barge-involving events (in particular, events of a barge passing through that lock). Consider now the singular proposition *that John loves John*. The claim that John occurs twice in the proposition *that John loves John* is even less problematic. Even on its face it parades as a count of events. With a suitably broad conception of events we can identify two such events: the event of John, the individual, being the agent of the proposition, and the event of that very same individual being the patient of the proposition. Same individual, different events. There seems to be no deep obstacle to extending this model to the case at hand: the two occurrences of the morpheme 'anti-' are two events involving one and the same morpheme.

(iii) Even supposing one can make good on these ideas, a number of residual mereological queries remain. Is there anything that overlaps the word 'anti-anti-missile' (where x overlaps y iff x and y share a part) but which does not overlap 'anti-missile'? If you answer 'no', then you have to depart from standard mereology, which tells us that if x is not y, then there is something that overlaps one but not the other. If you answer 'yes', then you have to find exotic objects that are suitable mereological difference makers. These decisions—intimately related to David Lewis's suspicions about 'structural universals'—are difficult, but they are not special to the mereology of words. They apply with equal force to other examples involving pairs of objects which are, roughly speaking, alike in matter but not form. Resolution of these questions is obviously beyond the scope of this paper.

principles. Second, we identify two of his central ideas that, even assuming those framing principles, strike us as seriously problematic, and we suggest how they can be improved upon.

One of Kaplan's guiding insights concerns the evolution of words: over time there may be considerable variation in how a word is written or pronounced. As it is transmitted from one epoch to another, it can change not only its phonological and orthographic contours, but its semantic and grammatical ones as well. Like Kaplan, we are after a metaphysics of words that satisfies this *Evolutionary Constraint.*

A second idea that is certainly implicit in Kaplan's account is that a word can be articulated in untold different systems. One and the same word can be written on a pad with a pen, typed on a sheet of paper, projected on a screen, spoken out loud, signed with a gesture, and Brailled on a plaque. Different media can be used to express the same message. Only our imaginations limit how we go about articulating words. Call this the *Multiplicity Constraint.*[33]

A third idea we shall try to respect is that two utterances or inscriptions, even within a particular community, may in an important sense of the word 'word' count as utterances or inscriptions of different words even though they are exactly alike in how they sound (in the case of two utterances) or in how they are written (in the case of inscriptions). Call this the *Coincidence Constraint.*

In addition to these three constraints, two further themes are central to Kaplan's conception of words and word performances. The first is his commitment to the importance of the concept of *repetition.* ("This notion of repetition is central to my conception" (103).) Suppose someone articulates a word in a performance. What makes a subsequent performance a performance of that word? (The question can be posed—though need not be—within the stage-continuant model: suppose someone produces a stage of a word-continuant. What makes a subsequent performance a stage of the same continuant?[34]) In answering this question, Kaplan is responsive to the fact that there need not be an intrinsic match between the performance of a speaker and a prior performance he has encountered in order for both to be performances of the same word: there may be significant differences in shape and sound. (To reinforce this point, notice

[33] There may be (contingent) priority relations between certain pairs of articulation systems. Thus one might argue that the alphabet works by representing phonology, rendering the former parasitic on the latter. We shall not enter into these priority disputes here.

[34] "What is it that makes a particular output, the transmission of the same word as that carried by a particular input?" Kaplan, *op. cit.*, p. 102.

that a speaker may encounter a spoken performance and then go on to produce a written one.) As a remedy to any presumed significance of an intrinsic match, Kaplan stresses in its place the importance of *intentional repetition*: what is important is that a speaker intends to *repeat* a word. This suggests the following sufficient condition for two performances to be of the same common currency name, which we shall call *Repetition*:

> *Repetition*: If p_1 is a previously encountered performance of a word w and p_2 is performed in such a way that its utterer intends p_2 to be of the same word as p_1, that suffices to establish that p_2 is a performance of w.

This sufficient condition hardly vindicates the importance of repetition, since it is compatible with *Repetition* that word matching is almost always secured by a different mechanism.[35] Kaplan is evidently committed to the pervasiveness of the repetition mechanism. What is the cash value of the requirement that repetition be pervasive? Given that, to our ear, it is not natural to speak of repeating except where some performance is being repeated, an instance of the pervasiveness of repetitive intentions suggests the following thesis:

> *Pervasiveness*: For any noncreative performance p_1 of a word w, there has to be a previously encountered performance p_2 of w such that p_1 is produced with the intention of repeating p_2.

The further central theme to which Kaplan is apparently committed is implicit in *Repetition*. Roughly speaking, the theme is *intending it to be so makes it so*, which we shall call the *Constitutive Role of Intention*:[36]

> *Constitutive Role of Intention*: If someone intends to produce the same word w as that used in a particular performance, then whatever comes out of his mouth (or from his pen) is a performance of w.

[35] Ruth Millikan, *Language, Thought, and Other Biological Categories* (Cambridge: MIT, 1984), p. 72 also ties word identity to a history of uses when she says, "the Martian who due to historical accident utters what sounds just like the French 'Il pleut,' even though he may happen to mean just what the Frenchman typically means when he uses this sound, does not utter the French but rather the Martian word 'il.'" But she minimizes the role of intention, "the parrot that mimics tokens of the English word 'hello' tokens *the English word 'hello'* and the child who uncomprehendingly copies out 'il pleut' writes down *the French word 'il'*."

[36] "The identification a word uttered or inscribed with one heard or read is not a matter of resemblance between the two physical embodiments.... Rather it is...a matter of intention....We depend heavily on resemblance between utterances and inscriptions...in order to divine these critical intentions....We also take account of accent and idiolect and all the use clues to intention. It is the latter that decides the matter." Kaplan, *op. cit.*, p. 104.

Of course, a performance may be in some sense a bad performance of a word. But, crucially, it is never a *non*performance: good or bad, it is a performance of that word. As Kaplan says:

> He may not do it well, from the external point of view. But it *is* what he is doing….No matter how poor the subject's *imitative* ability…we can imagine circumstances in which we would say, "Yes, he is repeating that name; he is saying it in the best way that he can." (103–04)

> …the difference in sound or shape or spelling, can be just about as great as you would like. (101)

Neither the emphasis on repetition (as a sufficient and pervasive condition of word matching) nor on the associated *Constitutive Role of Intention* is well motivated in Kaplan's discussion. Some straightforward observations will suffice to clarify our reservations.

First, notice that it is a grotesque exaggeration to suppose that when we produce a word we typically have in mind any particular performances which we intend to repeat. Intuitively, it seems that what we do is *pick up* a word and then, later, engage in performances with the intention of using *that word*. The original performances that allowed us to pick up that word are long gone from our mind or confused with others. The point is hardly unique to words. Introduced to a song or a dance, we may produce either without an episodic memory of the performances that allowed us to pick them up. To remedy this defect, an intermediate stage in word performance needs more emphasis.

Suppose we encounter some performance(s) that introduces us to a word. We thereby form an idea *W* of the word. This idea might outlive any memory of the performances that initiated it. We then engage in a performance that is controlled by an intention that involves *W*. At this juncture, it would be manifestly incorrect to describe the intention behind our performance as an intention to repeat, since there is nothing we are trying to repeat, nor anyone we are trying to mimic.[37]

To reinforce this picture, reflect on the myriad ways in which we can be introduced to a word. Although we might be introduced to it by a performance, we might just as well be introduced to it by description: "There is a word in English spelled 'b' followed by 'a' followed by 'b' followed by 'y' that means the same as 'infant'." This allows us to pick

[37] In discussion it became clear that Kaplan never intended to subscribe to Pervasiveness as we describe it. Given this, it was certainly a strategic error for Kaplan to emphasize the importance of repetition. At any rate, the idea presented in the above paragraph renders the concept of repetition altogether peripheral: when we speak we are certainly not repeating ideas.

up the word and form an idea of it. Yet when we subsequently utter 'baby' we are obviously not trying to repeat a performance of it, since we were never exposed to any performance of it.

Our first concern, then, is that (even noncreative) performances of a word rarely be controlled by a repetitive intention, and so Pervasiveness is *in*correct.

Our second worry concerns the presumed *Constitutive Authority of Intentions*. To begin, reflect on songs. In this case too, we can ask whether a performance being controlled by an intention to perform a particular song *S* guarantees that it is a performance of *S*. We take it to be obvious that in the case of songs a certain level of *tolerance* is in play: we allow a performance to be of a song even though it is a *bad* performance of it. But there also are myriad cases where even though there is a controlling intention to perform a particular song *S*, the performance counts as a *non*performance rather than merely a bad one: *tolerance has limits*. (It goes without saying that the boundary is vague, but that of course is no argument against the existence of the distinction.)

Suppose one intends to perform the song "Hey Jude" but merely grunts because of a defect or pathology: then the performance is a *non*performance, not merely a bad one. Suppose you intend to perform "Hey Jude" but mistakenly identify its lyrics and music with those of "Let It Be." Then your performance is a *non*performance rather than merely a *bad* performance of "Hey Jude" (leaving open whether it is a performance of "Let It Be").

The analogy between words and songs is arguably not perfect. In the case of songs, there is a sharp disanalogy between the relation of the vocal rendition of the song, on the one hand, and what occurs on the written score sheet for the song, on the other. The former is a performance of the song, whereas the latter is its mere representation. Arguably there is no such contrast between the spoken and written word (*contra* Aristotle).[38] But we do think that a nonperformance/bad performance distinction of the sort we just sketched carries over to words.

[38] We only say 'arguably': one line to pursue is that there is a deep analogy even here. Note that certain Kaplanian ideas about words carry immediately over to songs: (i) whether a performance counts as a performance of a song may be constitutively related to the community to which a performer belongs. Suppose a Martian wrote a symphony that resembles, but is still markedly different from, Beethoven's Fifth. Intuitively, a human performance may count as a bad performance of Beethoven's Fifth rather than a perfect performance of the Martian symphony owing to certain sociohistorical facts. (ii) Just as performances of a word may evolve in such a way as to change their character, so too with a song. As a song enjoys different renditions across time, its performances may change their character significantly. We return to these analogies in section IV.

If you intend to say the word 'dog' but instead merely grunt, then that is a nonperformance of the word rather than a bad performance. If Spooner intends to say, "The Lord is a loving shepherd," and instead says, "The Lord is a shoving leopard," then that is a nonperformance rather than a bad performance of what he intended to say. It is not that Spooner succeeds in his intention to say the word 'loving' but says it defectively (as when, for example, one mumbles or stutters); it is rather that he does not say the word 'loving' at all. If you see the word 'dog' for the first time and are ordered to re-write it but proceed to write 'g' followed by 'o' followed by 'd', then you have failed to write the word you were asked to: your intentions, as a matter of fact, lack constitutive authority. In effect, we are advocating *Tolerance*:

> *Tolerance*: Performance p is of a word w only if it meets relevant performance standards.[39]

Tolerance is anathema to Kaplan presumably because he thinks it bucks the Evolutionary Constraint. After all, evolutionary considerations allow that a word w pronounced 'dog' in one epoch is pronounced 'god' in a later one. Isn't this evolutionary insight incompatible with our claim that you have failed to say the word 'dog' if you produce something with 'g' followed by an 'o' followed by a 'd'? We presume not, so long as care is taken to realize that whether a performance counts as tolerable does not supervene on its intrinsic features but is instead determined by certain relational facts. Let us explain.

Assume, with Kaplan, that a word may evolve in its performance profile. Further suppose that the limits of toleration evolve as well: the line between nonperformance and bad performance itself may shift over time. Thus, for example, it is possible that there is a word w and two communities C_1 and C_2 such that the string 'god' is a tolerable way of writing down w in C_1, but not in C_2. These evolutionary

[39] Standards are typically generous. There are cases where the community can be brought to see handwriting as inscribing certain words even though this is not initially apparent. Similarly, there are cases where a community can be brought to see that someone is performing a certain sonata even though this is not evident, given that the performance is so bad. But still, tolerance has its limits. If someone, owing to a visual problem, sees a 'B' as a 'P', intends to repeat someone who says 'Big', and writes 'Pig,' then he fails to repeat the word he intended to repeat. In a footnote, Kaplan recognizes exceptions to the constitutive authority of intention, namely, cases where "*to the astonishment of the speaker* the wrong word came out….Some dark force has reached into the speaker's psyche and misdirected the hand of intention." Kaplan, *op. cit.*, p. 105n11. But notice that the 'Pig'/'Big' case is not quite like that. The speaker in Kaplan's case reacts with astonishment just by virtue of having witnessed his or her own performance. Not so in the 'Pig'/'Big' case.

considerations block sweeping generalizations such as "Nothing that looks like 'god' will ever count as a performance of the word 'dog'."[40] But this is perfectly compatible with the claim that if someone *in our community right now* writes down 'g' followed by 'o' followed by 'd', he will *not* have written the word 'dog' (even though he may have intended to do so).

To accommodate the evolutionary insight, we need to acknowledge that the limits of toleration are *local*. What is a tolerable articulation of 'dog' will be local to a particular community. Sometimes we are more tolerant of speech defects, foreign accents, and unusual scrawls than at other times. Shifts happen. In general, a necessary condition on articulating a word is that the articulation passes the standards of the local community.[41] We call this *Tolerance**.

> *Tolerance**: Performance *p* is of a word *w* only if *p* meets relevant *local* performance standards on *w*.

In this way, the evolutionary insight can be harmonized with the thought that there are limits in toleration of the sort we have gestured at. Since this harmony can be achieved, we see little advantage to the more radical tack that Kaplan is proposing.[42]

In sum, Kaplan's vision could be improved upon in two ways. First, he should have allocated a less central role to repetition. This need not mean we should ignore historical aspects of words. It is rather merely to allocate a suitably central role to the mediating ideas of words—ideas that may outlive memories of particular performances. Moreover, even if the object of those ideas is socially determined, it is going too far to *require* that we must encounter performances of a word in a community as a precondition of forming an idea of that word. The second way in which Kaplan's vision can be improved upon

[40] Furthermore, not just anything that looks like 'dog' counts as a performance of the word 'dog'. In this regard, though we agree with Herman Cappelen that intentions play a lesser role than Kaplan affords them, we disagree with him about whether the relationship between shape and expression is context *in*sensitive. See Cappelen, "Intentions in Words," *Noûs*, XXXIII, 1 (1999): 92–102.

[41] It is more accurate to require that one passes the standards of the relevant community, where in general, but not always, the relevant community is local. If someone is doing a performance in an attempt to simulate a fourteenth-century English speaker, then the relevant community will *not* be the physically local one. And of course there need not be pressure on an American speaker to produce the word 'schedule' in the British manner even if he is sitting at High Table. Indeed, it is preferable that he *not* try to mimic the local customs.

[42] It is also obvious that the tolerance conditions will be articulation-system dependent. Within a particular community there will be conditions on whether a vocal performance counts as an articulation of the word 'dog', and rather different conditions on whether a written or signed performance counts as one.

is to recognize that he overstates the constitutive importance of intentions. In order for a performance to count as a performance of a word, it needs to fall within locally acceptable limits. Not anything goes.

III. COMMON CURRENCY NAMES

The relevant sense of 'word' that forms Kaplan's putative focus—one that he takes to satisfy all three constraints from the last section and to be urged by them—is supposed to be one according to which two naming practices involving the string 'Peter' turn out to articulate two different names—that is, two different common currency names. There is an obvious sense in which a single name is involved here. In deference to this sense, Kaplan claims that there is a single 'generic name' in play. But he also claims that in such a case there will be a proliferation of 'common currency names', and in particular there will be a distinct common currency name associated with each distinct naming practice that uses 'Peter'. Indeed, it becomes clear that an important motivation for Kaplan's discussion of words is the demarcation and vindication of the category of common currency names.

For our part, we do think that there is one or more important sense of 'word' that answers to the Evolutionary, Multiplicity, and Coincidence constraints. But we harbor considerable skepticism about the category of common currency names. It is the project of this section to expand upon and justify that skepticism.

Kaplan makes it clear that he does not intend common currency names to be the designata of quote names—say, 'Peter'—as used in English, or at least as in philosopher's English. Certainly, such a theory would make a hash of the data, including the fact that it is perfectly acceptable to say, "They are both called 'Peter'." Rather, his idea is that there is a natural kind worthy of study which is important to philosophy and linguistics, and which answers to a way of theorizing about words that is prevalent in foundational philosophy of language and linguistics. According to this way of theorizing, various utterances, for example, of 'Paderewski' by English speakers (recalling Kripke's famous example—see Motivation Five below), may not all be utterances of the same name.[43]

The relation between the ordinary conception of names and the theoretical conception of 'common currency names' is delicate. We mention three options.

First, one might hold that in typical contexts the extension of 'word' in ordinary language is the set that includes common currency

[43] See Mark Richard, *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them* (New York: Cambridge, 1990); Richard K. Larson and Peter Ludlow, "Interpreted Logical Forms," *Synthese*, XCV, 3 (June 1993): 305–55; Fiengo and May, *op. cit.*

names, explaining away contrary data by claiming that ordinary people are prone to radical errors in judgments about words. For example, when ordinary people say things of the form "Two of us have the same name," or "That name can be used to refer to several different people," they typically express falsehoods. We see this position as something like a last resort. (To be clear, we see no evidence that Kaplan holds such a position.)

Second, one might hold that the extension of 'word' in ordinary use is never, or is hardly ever, a set that includes common currency names. On this conception, common currency names are irrelevant to ordinary 'word' talk. So, for example, even when people say "You and I have the same name," they may be speaking the truth because they are *not* expressing a proposition that requires a shared common currency name for its truth.

Certainly, Kaplan allows that in many contexts we are not talking about common currency words when we deploy the word 'word'. As noted above, Kaplan admits the notion of a 'generic name' and allows that two people can share the same 'generic name'. It is clear, in context, that he is thinking that when people say "We have the same name," they may be expressing the truth that people share the same generic name.

Yet it is hard to construe him as merely positing a class of entities—common currency names—that is theoretically important for linguistics, philosophy of language, and cognitive science, but of which ordinary folk have no conception. After all, much of what he says fails to square with this. For example, he frequently speaks of people intending to repeat a common currency name. But how can someone intend to repeat a common currency name if she has no ability to have singular thoughts about particular common currency names and has no concepts whose extension is the set of common currency names? There is nothing untoward about introducing a class of theoretical entities beyond people's ken. But it would be very odd at the same time to accord ordinary intentions towards those entities a crucial role in one's account of their place in the world.

On a third approach, ordinary people are *sensitive* to the presence of common currency names, and in some contexts they use 'word' to pick out a class that includes common currency names.[44,45] On this

[44] Our taxonomy is not exhaustive. For example, here is a fourth position which we find less appealing. There are such entities as common currency words; ordinary people are aware of their existence and think about them; but ordinary people never use 'word' to talk about them.

[45] Note, at the other pole, that there are contexts in which we use 'word' in a way that excludes both common currency names *and* generic names. Thus, for example,

approach we are at some level already accustomed to distinguishing different common currency names even if there are plenty of contexts in which our use of 'word' is insensitive to such distinctions.

On the first version of the project, one would expect there to be ordinary judgments about the sameness or difference of common currency names but should be very skeptical about their reliability. On the second version, one would expect there to be no ordinary intuitive judgments as to whether common currency names are the same or different. On the third version, such expectations would not obviously be out of place; nor would one expect them to be obviously unreliable. Kaplan's project is clearly of the third sort. His pronouncements make little sense otherwise. We shall continue the pursuit of common currency words in that spirit.

As we have already noted, it is central to Kaplan's vision that different common currency names may be vocalized by the string 'Peter' when those vocalizations belong to two different naming practices. Thus, for example, we use a different common currency first name for John Donne than we do for John Travolta, even though each name gets articulated by similar sounding performances. (That is not to say there might not be cases where we confuse these two individuals, and thus, where there is a single common currency name that we attempt to use both to refer to John Donne and to John Travolta. Suppose, for example, many different individuals were involved in the so-called Jack the Ripper crimes.) In addition, as already noted, Kaplan recognizes 'generic names', where John Donne and John Travolta, while they have different common currency first names, share the same generic first name. But if we take ordinary language data at face value, we automatically ought to be suspicious of common currency first names. We invariably speak of John Donne and John Travolta as sharing a first name, and it is rather difficult to access a reading of 'John Donne and John Travolta have different first names' under which it comes out true. Something needs to be done, therefore, to justify positing common currency names as corresponding to one important conception of names. Something needs to be done to motivate Kaplan's claim that "for serious semantics…it is my common currency conception that [is] important" (111). In what follows, we canvass and evaluate five such motivations.

---

in contexts where we ask how many words someone knows, both generic names and common currency names get excluded. Perhaps in part influenced by such ordinary usage, one or two linguists whom we have encountered claim that names are not words at all. However most linguists we have encountered doubt that there is any deep insight encoded by such remarks.

Motivation One: Words and Mentalese. Suppose, as many presume, that thought is conducted in an internal language—an 'I-language' or 'Mentalese'. It is eminently natural to hold that insofar as we distinguish two people in thought, we employ two different Mentalese tags for them. (One motivating thought is that we could not properly keep track of the sameness and difference of our own thoughts if there was pervasive semantic context-dependence even at the level of Mentalese.) Given this eminently natural picture, shouldn't we say that the public language performances of their names are articulations of two different names even if they sound alike?

Reply: This motivation plays illicitly on the verb 'articulate'. Suppose someone sometimes speaks French and sometimes speaks English. There may be some Mentalese expression for *rain* that he sometimes articulates using 'rains' and sometimes articulates using 'pleut'. But that does not mean that 'rains' and 'pleut' are the same word. There can be distinct words corresponding to the same Mentalese word.

Meanwhile, one might have different Mentalese vehicles in play when one uses 'red' to mean that something is red on the outside and when one uses 'it' to mean that something is red throughout. But this hardly means two English words are in play. Likewise, a Spanish-to-English translator may "articulate" one Spanish word by different English words depending on context. But that hardly means there are many Spanish words being translated.[46]

Motivation Two: The Intuitiveness of Kaplan's Three Constraints. Does the fact that there is something very intuitive about the ideas behind the Evolutionary, Multiplicity, and Coincidence Constraints establish that we already have a pre-theoretic conception of common currency names?

Reply: As Kaplan himself acknowledges, *even what he calls 'generic words' share this tri-fold profile* (115–16). The generic name 'John' was pronounced very differently in Shakespearean times than it is today;

[46] It bears emphasizing that while an individual might have two Mentalese words MW_1 and MW_2 for a particular pair of individuals *X* and *Y*, which he expresses using the same 'generic word' of English, he may *not* have two Mentalese names for those Mentalese words. Hence, he may not be capable of intending to use one of those Mentalese names rather than another. Thus, supposing MW_3 is the Mentalese name for the generic name 'John', he may be able to have the thought that the (generic) name 'John' is sometimes used as a name for *X*, but not that the Mentalese name MW_1 is a name for *Y*. On a related point, while he may be able to intend to use 'John' as a name for *X* on a particular occasion, he may be unable to intend to use 'John' to translate MW_1 on a particular occasion. We have already seen that Kaplan needs common currency words to figure in the intentions of language users. Even assuming multiplicity of Mentalese names, they may be unsuitable to ground the presumed multiplicity of word-theoretic intentions.

it also can be realized in many different systems of articulation—it can be spoken, written down, signed, Brailled, semaphored, and so on; the names 'Shawn' and 'Sean' are pronounced the same, though there are obviously contexts in which one ordinarily can count them as two names. In short, the intuitiveness of Kaplan's three ideas need not reveal any tacit commitment to an ontology of common currency names nor, in particular, to the thesis that in some ordinary sense of 'word' there are many different words that English speakers express by 'John', individuated by naming practices.

Motivation Three: Common Currency Names Are Needed for Kripkean Theories of Reference. Consider the following, fairly standard, simplified Kripkean model of reference:[47]

Stage One: A speaker associates a referent with a name by either a reference-fixing description ("Let 'Johnny' name the tallest male model in Alaska.") or by ostensive baptism ("We christen this ship 'Johnny'.").

Stage Two: That name is passed along the community by a causal chain of communication (transference). The reference of later uses of that name is determined by a historical chain that leads back to the reference-introducing event. The fundamental rule is this: the reference of a name (insofar as it has a reference) is determined by the Stage One event that associates some object with that name. Meanwhile, an appropriate causal chain renders some performance by an individual a performance of the same name that was introduced by a Stage One event. In short, a performance p_2 is of a name n only if there is an appropriate historical-causal chain connecting p_2 and some reference introducing performance p_1 of n.

As Kaplan is himself aware, this model requires a conception of names that aligns with the common currency conception (93–95).[48,49] There is a many-to-one relation between 'generic names' and Stage One events. If the same generic name 'John' is associated with John Donne and John Travolta, then it makes no sense to speak of the event whereby the reference of the generic name 'John' was fixed.

[47] Saul Kripke, *Naming and Necessity* (Cambridge: Harvard, 1980); Michael Devitt, *Designation* (New York: Columbia, 1981); and Richard, *op. cit.*

[48] Richard writes, "Part of what makes for sameness of name in the requisite sense is being part of the same 'causal chain' of transmission." *Ibid.*, p. 183.

[49] Kripke (*op. cit.*, note 9) writes, "…two totally distinct 'historical chains' that by sheer accident assign phonetically the same name to the same man should probably count as creating distinct names despite the identity of the referents." His commitments on the present topic turn on how 'totally distinct' is to be read. Is the use of 'John' to talk about John Donne '*totally* distinct' from the use of 'John' to talk about John Travolta?

For on the generic conception, 'John' can be used to refer to a variety of different people depending on context.

Reply: Granted, one cannot take on board the simplified Kripkean model in all its aspects and simultaneously hope that generic names will play the role allotted to names within the model. Yet one wonders how damaging it would be to alter the Kripkean model slightly, so as to eliminate the need for an ontology of common currency names. By our lights, what is crucial to the 'chain' idea is that the reference of particular *uses* of a name is typically deferentially determined. Suppose, for example, you encounter a particular use of the generic name 'John' by some utterer *U*. You may then use that generic name where your controlling intention is that your *use* of 'John' refer to whichever object *U* was referring to with his use of 'John'. In this way, the reference of the generic name 'John', on your occasion of use, is inherited from the reference of that name on *U*'s occasion of use. The central inheritance idea is thereby preserved *without* relying on an identity of a common currency name.[50]

Notice that this kind of model (unlike Kaplan's Kripkean one) can also readily account for cases where you quite obviously *inherit* reference without using the same generic name. Thus, suppose we hear someone talking about a guy called 'John', and he attributes some nasty acts and traits to him. One of us may decide to start calling him 'Mr. Nasty'. Here the reference has nothing to do with the preservation of a name. What is crucial is that the controlling intention is to use 'Mr. Nasty' in the relevant context to refer to whoever, in the context of conversation, had been referred to by 'John'.

In short, while one Kripkean theory of reference requires serious reliance upon an ontology of common currency names, it is not clear that any important insights would be lost by moving to a variant model that eschews any such reliance.

Motivation Four: Ambiguity and Polysemy. Linguists standardly distinguish ambiguity—as exemplified by the unrelated meanings for 'bank'—from polysemy—as exemplified by the family of related meanings for 'healthy' (as applied to animate objects, animal excretions, and food) and 'window' (as applied to both the physical object and its frame, as in "He opened the window" and "He went through

[50] There are slightly more complex models. For example, one might forge a Mentalese tag *T* for the object picked out by some use of a name *N*, and then use *T* to specify the intended referent of some later use of *N*. (Our wariness of the pervasiveness of repetition may encourage moving in the direction of some such model.) This model allows that some name *N* may on varying occasions have its reference fixed by differing Mentalese tags. But to infer that this implies that differing names are used on those occasions would be to commit the fallacy identified under Motivation One.

the window"). In the case of ambiguity, we would not be surprised to find two utterly different words in other communities, but in the case of polysemy we would. A standard piece of linguistic lore is that with ambiguity two words are in play but with polysemy only one with a variety of connected semantic potentials. In order not to prejudge certain questions, we will employ the technical term 'lexeme' for posits of linguist orthodoxy.

Various considerations induce linguists (and philosophers[51]) to say two lexemes are associated with 'bank'. One concerns distinct etymologies. It generally is taken to be a condition of ambiguity that the lexemes in question are known to have developed from what were formally distinct lexemes at some earlier stage: the use of 'bank' that means *a raised shelf or ridge of ground* has its roots in the Old Norse word 'banke', and its use that means *financial institution* has its roots in the Old French word 'banc'. Another consideration includes distinct distributional reflexes. For example, by and large, anaphora and ellipsis work differently for ambiguity and polysemy. We say, "He opened the window and went through *it*" (where 'it' is anaphorically tied to 'window' even though the two occurrences have slightly different meanings—physical object versus frame); similarly, we say, "He is healthy and *so is* the food he prepares for his family," but we do not say (except as a pun), *"He put some money in a bank and then swam to *one*," or *"After losing forty pounds, he is light and *so is* the color of his hair."[52]

Positing lexemes offers a potentially attractive theoretical explanation of these contrasts: an anaphor can be tied to the original only if the lexeme in play in the original is appropriate to the environment in which the anaphor appears.[53]

Mightn't we appeal to lexemes in vindicating this aspect of Kaplan's perspective?

[51] "If one of the tokens [of 'bank'] refers to a financial institution and the other to the edge of a river, it is implausible to insist that they belong to the same type." Zoltán Gendler Szabó, "Expressions and Their Representations," *The Philosophical Quarterly*, XLIX, 195 (April 1999): 145–63, at p.148.

[52] As a heuristic, linguists sometimes appeal to dictionary entries. While one expects multiple distinct dictionary entries for 'bank', one would not for 'red', 'window', or 'healthy'.

[53] It is not clear that the historical and distributional ideas are really complementary. Consider 'right'. It is standardly treated as ambiguous—largely for distributional reasons. But the uses of 'right' to mean *a direction* and to mean *correctness* have a common etymological origin. In the end, historical evidence plays an evidential, not a constitutive role in the thinking of most contemporary linguists on the subject: if 'right' has two distinct 'analyses' flowing from distinct 'lexical entries', then there is more than one lexeme in play even if there is a unity in historical origin for both analyses.

Reply: We have already acknowledged that there is flexibility in which objects we talk about in using 'word'.[54] Sometimes, for example, we treat quotation words as picking out something that can be either written or spoken, but other times as picking out the very sign used to articulate words. Thus, for example, we might in one context agree that 'red' is the same word as 'RED' (using a more abstract sense of 'word'), while in another context agree (using quote words to name the signs themselves) that 'red' is in lower case, while 'RED' is not. (Consider similarly, "'red' occurs three times in the second paragraph, but 'RED' only once.")[55]

We concede further that there are uses of 'word' that seem to point in the direction of lexemes. The relevant frame of mind here is one that Lyons invoked in his seminal discussion of lexemes,[56] one where we are willing to make such claims as:

(1) 'Find' and 'found' are versions of the same word.

Claims like (1) abstract away from the surface in a quite radical way. In that frame of mind, it is natural to say (2):

(2) 'Will' and 'willed' are versions of the same word.

In that frame of mind, it is also natural to say (3):

(3) 'Will' and 'would' are versions of the same word.

But it is not natural to say that 'willed' and 'would' are versions of the same word. In this frame of mind, then, it would seem that we are committed to there being *two* words associated with 'will'. Let us grant, then, that there are contexts in which 'word' is being used to pick out lexemes.

Crucially, however, it is very far from clear that this can serve as a basis for vindicating Kaplanian common currency names. To do so would require that we say that two different lexemes are manifested by 'John' as it occurs in 'John Donne' and 'John Travolta'. But what justifies this? The fact that different people are denoted in the various contexts of use does not settle the issue—after all, polysemous uses of a single lexeme generate varying denotations as well. Insofar as the data provide guidance, they do not seem to support an ambiguity

[54] In the language just introduced, perhaps the right conclusion is that 'word' is itself polysemous.

[55] For an explanation of the phenomena that does not invoke ambiguity or context sensitivity, compare with Cappelen and Ernest Lepore, *Language Turned on Itself: The Semantics and Pragmatics of Metalinguistic Discourse* (New York: Oxford, 2007).

[56] Lyons, *Semantics: Volume 2*.

thesis—after all, it is felicitous to say, "She is Janet and [pointing to a different person] so is she," and "Every Janet that I know…." The felicity of such statements tells against ambiguity and hence lexeme proliferation.[57] The claim that there are many lexemes associated with the use of 'John' in English is simply bad linguistics. Insofar as names are identified with lexemes, that tells against Kaplan's common currency views, *not* in favor of them.

Motivation Five: Attitude Reports. Some accounts of attitude reports rely heavily on lexical considerations, treating some attributions as true only if the speaker bears a certain relationship to words in a report's complement clause.[58] In these contexts, it is alleged that form trumps content and reference. (4) and (5) attribute distinct beliefs to Lois because they relate her to different names.

(4) Lois Lane believes that Superman can fly.
(5) Lois Lane believes that Clark Kent can fly.

Insofar as one goes in for this type of account, there is pressure towards admitting that a common currency conception of names is central to ordinary practices of belief reporting. After all, this account is supposed to extend to Kripke's well-known 'Paderewski' case.[59]

Peter is given two bits of testimony from two different people: "Paderewski has musical talent," and "Paderewski does not have musical talent." Peter does not realize (let us suppose for a good reason) that each piece of testimony is about the same person and so accepts both pieces of testimony. We look on. We are tempted to say, "Peter believes that Paderewski has musical talent," "Peter believes that Paderewski has not musical talent," and that Peter is not being irrational. The solution proposed by the relevant accounts is that the pair of reports relates Peter to distinct names. But this apparently requires that there are two different names in play and that this fact is tacitly recognized by our reporting practices. Thus,

[57] We again note in passing that some contemporary linguists refuse to treats names as words, but we put that issue to one side.

[58] For example, Gilbert Harman, "Logical Form," *Foundations of Language*, IX, 1 (September 1972): 38–65; James Higginbotham, "Linguistic Theory and Davidson's Program in Semantics," in Lepore, ed., *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Malden, MA: Blackwell, 1986), pp. 29–48; Higginbotham, "Belief and Logical Form," *Mind and Language*, VI, 4, (Winter 1991): 344–69); Kaplan, *op. cit.*; Larson and Ludlow *op. cit.*; Larson and Segal, *Knowledge of Meaning* (Cambridge: MIT, 1995); Richard, *op. cit.*; and Segal, "A Preference for Sense and Reference," this JOURNAL, LXXXVI, 2 (February 1989): 73–89.

[59] Kripke, "A Puzzle about Belief," in Avishai Margalit, ed., *Meaning and Use* (Boston: D. Reidel, 1979), pp. 239–83.

Ludlow and Larson infer that these homophonous expressions "are actually two names here, 'Paderewski$_{\mathrm{I}}$' and 'Paderewski$_{\mathrm{II}}$'."[60]

Reply: Note first that if this style of account is right for identity confusion, then one ought to expect it to apply to a range of cases where proper names are not in play. For example, suppose someone is told, "No one has ever lived here," and then later told, "Someone once lived here," and while he has excellent evidence that 'here' refers to different places, in fact it refers to the same place twice over. This has a very similar structure to the Paderewski case, and surely demands a similar solution. But the solution in question, as applied to this case, requires us to say that two words are in play, 'here$_1$' and 'here$_2$', and not one. Certainly, some will be unwilling to absorb this consequence—but it is important to see that it needs to be absorbed.

Our main objection to the account in question is that it flies in the face of something that is right about Kaplan's account. We do not think merely intending to use a word one has picked up guarantees success. However, if one satisfies local performance standards, it is hard to deny that intending to use a word will bring success in its wake. Now, suppose someone at one time tells us "Paderewski is musical," and at another time, "Paderewski is not musical." We repeat back, "Paderewski is musical," and "Paderewski is not musical."

[60] See Ludlow and Larson, *op. cit.*, p. 319. A related but possibly distinct motivation focuses on logical truth itself. We say 'Hesperus is Hesperus' is a logical truth but not 'Hesperus is Phosphorus'. This differential judgment appears to rely on the recognition that 'Hesperus' and 'Phosphorus' are distinct names. But suppose now that there are two different naming practices wherein the generic name 'John' is associated with an individual. (Suppose John Lennon and John Travolta are in reality one object with two personae. Isn't it tempting to think that 'John is John' is not a logical truth even though the same generic name appears on each side? If name individuation is to provide the basis for an account of logical truth, it seems that generic names cannot do the work. Considerations such as these are evidently important to Kaplan's own motivations. See, for example, Kaplan, *op. cit.*, p. 94. Now Kaplan is sensitive to the fact that even common currency name identity in '$a = a$' will not secure a sense of triviality in every case given that one might store the same common currency name twice over without realizing it. But one might hold that in such cases there is a logical truth that the speaker fails to recognize. By our lights, an appeal to common currency names will only be appealing at the foundations of logic if one is willing to make similar appeals to a fine-grained ontology of the words 'here' and 'that', appeals about which we are skeptical. (The point extends to predicates: for any context-dependent predicate, *F*, a vindication of the claim that 'All *F*s are *F*' is true by virtue of its form will have to rely on a fine-grained ontology of predicates.) Further, we note that given that common currency names arguably can switch reference, it is a short step to allowing that there are cases where some uses of a single common currency name refer to one individual, others to another, and hence, cases where an identity claim flanked by two occurrences of the same common currency name is false. The fact that even common currency names arguably can have variable referents entails that even they cannot, after all, very well play the role of the individual constants of logic. This all raises important issues about the form-theoretic conception of logical truth that we cannot pursue further here.

Suppose we have good evidence she is speaking about different people but is trying to confuse us. It is clear enough that the speaker is using the same word in both performances. It is also clear enough that on each occasion we intend to repeat the name we are confronted with and that we satisfy local performance standards for that name. Given all this, it is hard to deny that we use the same word as the speaker on each occasion of repetition. Given symmetry and transitivity of identity, it follows that we both use the same name on each occasion. But this is hard to square with an account of the relevant belief ascription that depends upon our using *different* names on each occasion.

We do acknowledge that something in the vicinity of these accounts nevertheless may be correct. Ludlow offers an account of the Paderewski case according to which we, the ascribers, make up two names in some internal language 'on the fly' with which we label the Paderewski nodes of the semantic trees. Our sense that we are not convicting the subject of incoherence is then explained by the fact that we use different labels in each of the ascriptions.[61] Here is not the place to evaluate this account, which relies on labeling opportunism. But note, crucially, that this account does not require that the subject of the belief ascription use two different names in a Paderewski case. We suspect that insofar as there is anything to the idea that Paderewski cases are to be explained by associating the subject with two different labels, it will not rely on an ontology of common currency names.

In sum, we have yet to see a powerful case for the thesis that in some ordinary contexts, 'word' has Kaplanian common currency names in its extension. Perhaps a fine-grained ontology that associates multiple lexemes with 'bank' can be justified on theoretical grounds, and perhaps there are settings in which we use the word 'word' to talk about lexemes. But even so, the case remains to be made that Kaplanian common currency names are the objects of ordinary repetitive intentions or ordinary 'word' talk. Certainly, none of the justifications we have examined are particularly persuasive.

IV. CRITERIA OF IDENTITY

We have voiced multiple reservations about Kaplan's account, and along the way sketched an answer to the first of the pair of questions with which we began. But what of the second question?

Criteria of identity come cheap if we allow them to go trivial. To what extent can we hope to come by both true and informative criteria of

[61] Ludlow, "Interpreted Logical Forms, Belief Attribution, and the Dynamic Lexicon," in Katarzyna Jaszcolt, ed., *The Pragmatics of Propositional Attitude Reports* (New York: Elsevier Science, 2000), pp. 31–42.

the identity of words? Given the flexibility in our uses of 'word', the context of our inquiry needs to be made clearer before we press ahead. The preceding discussion indicates that the theoretically deepest—and of most interest to philosophers—notion of 'word' is one according to which it is used in a way that conforms to the lexeme conception. We ask the reader to understand us to be using 'word' in that way, and to read what follows in that light. We shall suggest that even so restricted an expectation of identity criteria for lexemes to a large extent is misplaced.[62]

On the lexeme construal, we must institute a sharp distinction between words and their performances. With that distinction in place, pursuit of a criterion of identity can take either of two forms. (Here we borrow some terminology from Williamson.[63])

First, we may attempt to articulate interesting *first-level identity criteria*, where the objects for which the criterion of identity is stated are the same as those between which the criterial relation obtains. The criterion of identity for sets stated by the Axiom of Extensionality (sets are identical iff they have the same members) and Davidson's criterion of event identity (events are the same iff they have the same causes and effects) are examples of first-level identity criteria.

Second, we might pursue a *second-level identity criterion*, where we state a criterion for two objects from a different domain to stand in some particular relation of interest to the same object from our original domain of interest. The Fregean criteria of identity for directions and numbers are both second level. Two lines (neither of which is itself a direction) are of (here 'of' marks the particular relation of interest) the same direction iff they are parallel; two classes are numbered by the same number (here 'numbered by' marks the particular relation of interest) iff there is a one-one map from one to the other.[64]

[62] Certain considerations that follow will carry over to uses of 'word' that are slightly less abstract than the lexeme use. We shall not, however, be asking whether and to what extent there are identity criteria corresponding to each use of 'word'.

[63] Williamson, *Identity and Discrimination* (Cambridge, MA: Blackwell, 1990).

[64] At either first- or second-level identity criteria, we might less ambitiously seek only (interesting) necessary or sufficient conditions. Thus, at the first level, we might seek interesting claims of the forms:

Words w_1 and w_2 are identical only if R holds between w_1 and w_2.
If R holds between words w_1 and w_2, then $w_1 = w_2$.

Meanwhile, at the second level we might look for interesting claims of the forms:

Performances p_1 and p_2 are of the same word w only if R holds between p_1 and p_2.
If R holds between performances p_1 and p_2, p_1 and p_2 are of the same word w.

Also, at both first- and second-level criteria of identity, we might pursue something more ambitious. Take the above second-level criterion for directions. It is an intra-world

Note that it should not strike anyone as particularly alarming if we fail to devise informative first-level criteria of identity. After all, the standard criteria of identity for directions, for example, are second level and *not* first level. (The prospects for informative first-level criteria of identity for directions would be made worse if we allowed for directions that are not the directions of any line. Assuming there are unarticulated words, the situation with words is akin to that bleak situation.) Nevertheless, let us consider an attempt at a first-level criterion for word identity from Richard and Millikan:

> *Origin*: Words w_1 and w_2 are identical iff w_1 and w_2 have the same originating event (where the originating event is the first performance of a word).

A few observations about Origin are in order. First, it assumes that the relation of 'performance of' between an event and a word is already (sufficiently) well understood. Second, it assumes there are no unperformed words. And third, it assumes there cannot be a tie as to which events are the first performances of a word. Origin's take-home message, in short, is that there is exactly one event that is the first performance of any word. (Origin is equivalent to that message, since the latter is obviously derivable from the former and, given Leibniz's law, the former is derivable from the latter.) As such, Origin is far less informative than one might initially think. For example, it is logically compatible with Origin that only one word exists: to derive the obviously correct conclusion that 'cat' is *not* the same word as the word 'dog' you need additional information about which was the first performance of each. In any case, if what we have said about unarticulated words is correct, at least one of the assumptions upon which this first-level criterion depends is *not* true.

The following first-level criterion *is* arguably correct:

> *Performance*: Words w_1 and w_2 are identical iff it is not possible that some performance is of w_1 but not of w_2.

Unfortunately, all that Performance encodes is that *if* there are two distinct words, then it is possible that there be a performance of one without being of the other, and this does *not* seem particularly interesting.

---

criterion; it tells us when two lines in the same world partake of the same direction. But it does not obviously tell us anything about when lines in different worlds partake of the same direction. One might not think that *being parallel to* is a relation that can obtain between things in different worlds.

Do second-level criteria of identity fare any better? In the case of directions, we have an adequate second-level criterion already in hand: lines share the same direction iff they are parallel. (Note that this second-level criterion would not be impugned even if there were directions that were not the direction of any line.) With words and their performances, the situation appears to be much less promising. We need to face the prospect that there is no easily statable relation between performances that provides an interesting condition that is also both necessary and sufficient for their being performances of the same word.

Let us begin by engaging with a more modest task, namely, that of providing interesting necessary *or* sufficient conditions.

A candidate sufficient condition, Kaplanian in spirit, is *Intention*:

> *Intention*: If performance p_1 is *intended* to be a performance of the same word w as performance p_2, then p_1 is a performance of the same word as p_2.

We have already seen that Intention fails: among other things, certain grunts and groans are not performances of words no matter how much we intend them to be so.

A candidate for an interesting necessary condition on word identity, also inspired by Kaplan, is *Connection*:

> *Connection*: If p_1 and p_2 are performances of the same word w, then there is a performance p_3 of w such that p_1 and p_2 are both *historically/causally connected* to p_3.

Connection also fails, and for reasons we have already provided: performances p_1 and p_2 might be built out of morphemes presented separately to each performer, with neither traceable back to a common cause that takes the form of a performance of the word w of which each is a performance.

We do *not* however wish to preclude interesting necessary conditions that are of philosophical interest. For example, a much weaker version of the causal idea to which we are somewhat more sympathetic is *Isolation*:

> *Isolation*: If linguistic communities c_1 and c_2 are causally isolated, then performances p_1 in c_1 and p_2 in c_2 are not performances of the same word.

Some have the intuition that a perfect intrinsic match among performances suffices for a word match.[65] But this intuition dissolves once

[65] See, for example, Cappelen, *op. cit.*, pp. 95–96.

one sees that the performance profile of a word can evolve.[66] Suppose two isolated communities have words pronounced as 'moan' and 'mean', respectively. There is no inclination to say that a single word is in play. Suppose the performance profile of 'mean' evolves so that its performances are 'moan'-like. We are tolerant enough evolutionarily to allow that the same word is in play in the latter and earlier performances. But the logic of identity forces us to admit that different words are in play in latter performances of 'mean' and current 'moan' performances of the other community, despite an intrinsic match. (Of course, it is obvious that no second-level criterion of identity can be wrought from Isolation. At best we get a weak, though mildly interesting, necessary condition for word matching.)

The task of providing an interesting necessary condition on word matching between performances did not prove hopeless. And our survey of candidate sufficient conditions has been far from exhaustive. Ought we to be pessimistic about the possibility of a satisfying second-level criterion?

In this connection, it is worth underscoring a particularly satisfying feature of the second-level criterion of directions: namely, the relation of being parallel is a relation we can grasp independently of an ontology of directions. Call a second-level criterion '*quasi-reductive*' if the relation that appears on the right-hand side of the criterion is capable of being grasped independently of the domain for which one is providing the second-level criterion. In the case of words, the hope for a quasi-reductive criterion seems rather dim. Any criterion that makes reference to intentions to repeat a word, to use the same word, and so on will be inappropriate to the quasi-reductive project. And we see no prima facie reason for optimism that the stock of relations compatible with the quasi-reductive project can provide the basis for a criterion that comes close to being compelling.[67] In short, the relations that are naturally appealed to when concocting a second-level criterion are relations that cannot underwrite a quasi-reductive criterion.

Of course, assuming both the existence of words and that the facts about word matching as between performances supervene on the microphysical ground floor, there presumably will be some extremely complicated second-level criterion that is extensionally correct, and some yet more complicated second-level criterion that will be necessarily

[66] It also dissolves once one gets into the lexeme-theoretic conception of words, according to which 'will' and 'would' are versions of the same word, 'will' and 'willed' are versions of the same word, but 'willed' and 'would' are not versions of the same word.

[67] It goes without saying that the hunt for interesting inter-world criteria will be even more problematic.

correct. However, this by no means implies that we have epistemic access to such criteria, nor that the success and well functioning of thought and talk about words relies on some tacit grasp of such criteria.

Now, from the armchair we have no general proof that compelling second-level criteria are unavailable. Certainly, the abject failure of attempts to provide criteria of this sort does not inspire hope that an interesting second-level criterion will be forthcoming. But in advocating pessimism about second-level criteria we do not merely wish to rely on a pessimistic induction. We shall, in what follows, sketch a picture that we believe is independently plausible and which suggests that the pursuit of second-level criteria is rather misplaced. We shall, along the way, ask a related, urgent question: why is an ontology of words-as-lexemes legitimate, even in the absence of a second-level criterion of identity?

To begin, it is useful to recognize that the situation with words is in many ways analogous to the situation with dances and games. In each of these cases, the following features are in play. First, community members have a pretty good appreciation of locally acceptable performance standards. That is to say, they have a pretty good capacity to distinguish performances from nonperformances in their local environment. Second, community members have an evolutionary conception of the relevant entities: dances may have been performed differently and games may have been played differently in earlier epochs. But, third, there is vast indecision about the distinction between cases where a new dance or game comes into existence and cases where it is merely an old game or dance in different clothing.[68]

Notice that an appeal to some crude causal criterion is of little use in resolving such indecision. Cases where a new dance is inspired by an old dance and cases where old and new performances are varying performances of the same dance are all cases of causal connectedness between the old and new.

As with dances, there will be vast indecision as to when words comes into existence. Consider, for example, 'moan' and 'mean'. Most English speakers take themselves to know that local performances of each are performances of different words, and hence, they take themselves to know that 'moan' and 'mean' are different words. Yet when confronted with the historical facts we have no clear judgment as to when each sprang into life.[69] There are causal lines

[68] It is sometimes said that all that holds various games from different locales and epochs together is their name. Obviously, this maneuver, whatever its other merits, will not suffice for individuating words.

[69] Cf. Cappelen, *op. cit.*, p. 95.

from both clusters of performances to a cluster of performances in Saxon England. We are less tempted to say that there were two words in play at that point. And yet we have no clear judgment as to when our words were first performed in the history of speaking peoples. Similarly, we have no clear judgment as to whether some distant causal predecessor of this or that word among Latin or Greek peoples was our word or merely an inspiration for it. This indecision will not disrupt actual practices very much—the discriminations we are called upon to make in practice are almost always local, and indecision in this domain is not nearly as rampant.

As we have already hinted, when an ontological posit has the kind of profile just gestured at, we can scarcely hope to defend its legitimacy on the model of the abstraction principles standardly invoked as foundational to the talk of directions and numbers. Even within our own locale, nothing cleanly marks word matching between performances in the way that parallelism marks direction matching between lines or how the existence of a one-one function marks number matching between pluralities. And the situation is worse still when we move beyond that locale, owing to the vast stretches of indecision.

It is common enough for the metaphysician to encounter the kind of profile just described. There are two standard reactions. The natural-kind optimist supposes there are natural contours in the world that favor certain candidates over others: while our knowledge of word matching may not favor one candidate (or one small cluster of candidates) over others, one may still count as the semantic value on account of its being more natural, joint-like, and less gerrymandered than the others. Hope might even be held out that empirical inquiry will uncover some such boundaries. Perhaps, for example, empirical inquiry about the structure of morphology or psycholinguistics will resolve questions of word matching one way rather than the other.

A second reaction—that of the gruesome pessimist—is the same as the first except that it does not assume underlying joints in nature that make it natural to resolve questions one way rather than another.[70] Instead, there is a plenitude of packages of objects and exemplification relations in the vicinity of talk about words, none more natural than the others, each fitting the extent of our knowledge about word matching. The candidates give different verdicts to questions about word matching.

The natural-kind optimist calls upon us to collaborate with lexicographers, linguists, psychologists, and anyone else who might be

[70] Note that the sloppy realist can still agree that the lexeme conception is in some interesting sense the theoretically deepest, since the latter point does not entail that anything like the claim that criteria of identity will be forthcoming.

able to marshal data. The gruesome pessimist advocates a benign quietism with respect to any departures from clear cases. There need not be a correct answer forthcoming, or else we are utterly unable to know what it is.

Now, the appearance of the category of lexemes in theoretical linguistics may seem to offer hope to the natural-kind optimist. But such hope may well be premature. When one looks at discussions of lexemes within theoretical inquiry, confident judgments of identity and difference—and associated diagnostics—are typically local. For example, the tests we have looked to for a single lexeme in the case of polysemy are designed to apply to a temporally and communicatively local set of uses. When it comes to performances that are temporally and/or spatially distal we get little systematic guidance. Perhaps there is a case to be made for natural-kind optimism, but most of the work remains to be done. The actual scenario may well be bleaker: there are no joints in nature that can resolve distal questions of word identity one way or another.

There are two subspecies of the gruesome pessimist. One is the sloppy realist. On this picture, the unsettled questions turn out to rest on borderline cases and are to be handled using the correct theory of vagueness (whether it be epistemicist, supervaluationist, or whatever). In that case, there either are facts we may never know or simply no facts at all about the myriad borderline cases left unresolved by our capacity to settle questions in the area.

Assuming that rampant semantic indecision does not mandate repudiation of this or that ontological posit, we are left with a practice that, while not so bad as to warrant skepticism, is too flimsy to warrant sustained metaphysical investigation. On this perspective, there is good news and bad news. The good news is that the elusiveness of questions of word individuation need not indict our practice of positing words. The bad news is that the accessible facts about words run so shallow that there is little philosophical payoff to ruminations about word identity. Those who pursue questions of word individuation and hope for systematic answers are almost invariably in the grip of a faulty picture of the semantic mechanisms that underlie thought and talk about words.

A second version—atheism—*denies* the existence of words. Words are a mistaken ontological projection. There are interesting relations of connectedness among performances, but no interesting equivalence relation. A failure to recognize this has led to an unwarranted postulation of objects whose coherence is ultimately undone by the lack of grounding equivalence relations. (It is as if we posited directions in a setting where it turned out that *being parallel to* is not an

equivalence relation). Note that the view of words as continuants does not fare any better here. If there are no interesting equivalence relations of connectedness, then there is no useful way of gathering performances into continuants either.[71]

The defense strategy for the sloppy realist is clear enough. She will emphasize that words are no worse off than songs, dances, and so on. Unless some general atheism about the posits of common sense can be motivated, then atheism in this case is likely to be an overreaction. Is there a productive way forward for the atheist? One promising way may be to focus on 'fission' cases. Suppose x belongs to a community that uses a particular word, 'happy'. Two communities c_1 and c_2 pass by x's community and, by x's lights, appear to pick up that word and return to their homelands with it. x has a description of this case that by her lights is extremely natural. c_1 picks up that word and comes to pronounce it in one way, while c_2 picks up the word but comes to pronounce it in a very different way—let us say 'harpy' and 'hapry'. Suppose c_1 and c_2 come to attach different meanings to the relevant words. Again, this will have a natural description by the lights of x: c_1 and c_2 use the same word with slightly different meanings. Suppose c_1 and c_2 encounter each other. The c_1 members treat the descendents of 'happy' in c_2 as articulations of a different word than the descendents of 'happy' in c_1. Similarly for c_2. But a natural perspective on the situation by the lights of x will be that c_1 users and c_2 users do not realize they are using the same word. Indeed, very natural considerations will *force* such a conclusion. After all, it is eminently natural to think that c_1 is using the same word as the original, albeit with a different meaning, and that the same is true of c_2. Given these facts, x will be forced to think that the two communities are using the same word without realizing it. But now think of things from the perspectives of c_1 and c_2. Suppose the c_1 users pick up the c_2's uses of 'hapry' and the c_2 users pick up the c_1 uses of 'harpy'. Members of both c_1 and c_2 will find it natural to think they are using two different words. Even if they learn that there is a common origin, this likely will not affect that judgment. The situation will be similar to that of the English 'moan' and 'mean', which seem obviously to be two words despite a common origin in Anglo-Saxon ancestry.

So who is right? x finds the claim 'they are both using the same word' to be obviously true. The c_1 and c_2 users find the claim 'we are both using different words' to be obviously true. Reckoning both

[71] This strategy adapts what Kit Fine has claimed about meanings to the level of words themselves. See Fine, *Semantic Relationism* (Malden, MA: Blackwell, 2009).

judgments to be borderline is very disturbing, for it seems to imply that our own judgment that 'moan' and 'mean' are different words is at best borderline. But siding with one perspective seems bizarrely chauvinistic. Atheism offers a way out of the dilemma: there can be a satisfying reconciliation if only they abandon the ideology of word identity for that of connectedness. All parties can agree that the c_1 uses are connected to the x uses, that the c_2 uses are connected to the x uses, but that the c_1 uses are not connected to the c_2 uses. But what reconciliation does the sloppy realist have to offer? She might look for guidance at what metaphysicians have to say about other fission cases. Where there is person fission, some metaphysicians say there were two people all along. But the judgment that there are two words all along in the case just described seems almost as disturbing to the natural conception as atheism.[72] Given the prima facie conflict with common sense, atheism admittedly faces an uphill battle. But it is far from clear that it is a battle that cannot be won.

In conclusion, let us turn briefly to the question with which we began this section. Given the polysemous nature of the word 'word', there is no context-independent answer to the question "What are words?" any more than there is a context-independent answer to the question "What are books?" Nevertheless, Kaplan was onto something. There are many contexts in which the word 'word' picks out neither particular performances nor an abstract entity which merely serves to encode the superficial form of particular performances.[73] That said, it is doubtful that the best treatment of such uses adopts a conception of words as four-dimensional continuants. Rather, there is more hope for a model of words as abstracta, though one that breaks with the standard type-token model's picture of the relevant abstracta as pattern-like. The challenge remains to provide a criterion for word matching between performances. Having eschewed superficial criteria, this challenge takes on a forbidding character, since there is no obvious surrogate that can provide a criterion once the form-theoretic ones have been dispensed with. Rather than expect that such criteria will be forthcoming, we must take seriously a conception of our practice which guardedly endorses on ontology of words while despairing of such criteria. Whether the reflective metaphysician ought

[72] One possible way out for sloppy realism is a sneaky contextualism. The c_1 and x users pick out a different domain by 'word'. Such maneuvers seem, *prima facie*, rather desperate.

[73] Indeed, the context of this sentence—where we use the expression "the word 'word'"—is one such context.

to sign on to such an ontology as well is a vexed question. But whether one opts for sloppy realism or atheism, one thing is clear. As far as individuation criteria are concerned, a lexeme-like conception of words, even when supplemented with the tools of theoretical linguistics, may have just the wrong mixture of superficiality and indecision.

JOHN HAWTHORNE

Oxford University

ERNEST LEPORE

Rutgers University