DOI: 10.1111/phpr.12781

ORIGINAL ARTICLE

Philosophy and Phenomenological Research

Just Words: Intentions, Tolerance and Lexical Selection*

Una Stojnić

Department of Philosophy, Princeton University

Correspondence
Una Stojnić, Department of Philosophy, Princeton University.
Email: ustojnic@princeton.edu

Abstract

We all make mistakes in pronunciation and spelling, but a common view is that there are limits beyond which a mistaken pronunciation or spelling becomes too dramatic to be recognized as of a particular word at all. These considerations have bolstered a family of accounts that invoke speaker intentions and standards for tolerance as determinants of which word, if any, an utterance tokens. I argue this is a mistake. Neither intentions nor standards of tolerance are necessary or sufficient (individually or jointly) for determining which word an utterance tokens. Instead, drawing in part on empirical research on word production, I offer an alternative account, Originalism-plus-Transfer (OPT), according to which word tokening depends entirely on lexical selection during word production, and on how the selected lexical item is situated within the network of causal/historical connections leading back to its neologizing. Once the elements of my account are in place, as a bonus, we will have resources for a promising answer to the question of word individuation as well.

*Acknowledgments: Earlier versions of this paper have been presented at a symposium at 2021 Meeting of the American Philosophical Association, Eastern Division, 2020 SLOW workshop, and 2019 Rutgers Alumni Semantics Workshop. I am grateful to the participants at these events for the feedback. I would also like to thank Louise Antony, Josh Armstrong, Michael Glanzberg, Zoltán Gendler Szabó, Jeff King, Kyle Landrum, Ernie Lepore, Jürgen Lipps, Eliot Michaelson, Paul Pietroski, Alex Radulescu, Achille Varzi, and an anonymous reviewer for *Philosophy and Phenomenological Research* for helpful comments and discussion.

1 | INTRODUCTION

Once I hear or read the word 'word,' when are my subsequent pronunciations or spellings tokenings of the same word? Two commonsensical observations inform our answers: first, speakers are not infallible, so there has to be some tolerance for error. I can still token 'word' even if I mispronounce it slightly or misspell it somewhat. But, second, not *anything* goes: a dramatic mispronunciation or misspelling will make it impossible to understand me. If I try to utter the word 'word,' yet what comes out is a grunt, haven't I failed altogether? These intuitions are typically explained by appeal to two types of consideration. First, speaker intentions matter: whether a speaker uttered, but possibly mispronounced, the word 'word,' depends on whether they intended to utter it. But second, if the deviation is so great so as to exceed the limits of tolerance, then even if the speaker intended to token a particular word, they will have failed. It is thus no surprise that most theorists emphasize speaker intentions and/or a threshold for tolerance for deviation from canonical articulation in explaining what it takes to token a particular word.[1]

It is commonly either implicitly or explicitly presumed that intentions are necessary and/or sufficient for tokening a word (Bromberger 1989; Kaplan 1990, 2011; *inter alia*).[2] To avoid the conclusion that anything goes, it is often added that an utterance has to satisfy a standard:[3] it must not deviate from its canonical articulation to a greater degree than specified by some (contextually determined) standard of tolerance within a local linguistic community (Hawthorne and Lepore, 2011).[4] These considerations, though natural, I shall argue are mistaken. Neither speaker intentions nor conformity to the standards of tolerance are necessary or sufficient (individually or jointly) for utterances to be of the same word. But what then is left to explain when they are? The aim of this discussion is to develop a positive account according to which whether utterances are of the same word depends both on the process of speech production, and the history of the speaker's word acquisition.

Before proceeding, it is worth pointing out that, despite the appearance of esotericism, the topic of when two utterances are of the same word is quite significant. We communicate with words by relying on our capacities to token the same words, and recognize which ones are tokened. If I want to convey a thought to you, I should utter words that (in this context, in a language we both understand)

[1]Talk of "canonical" articulations is an idealization: a word can have more than one "canonical" articulation, e.g., /ˈskɛdʒuːl/ vs. /ˈʃɛdjuːl/, or 'colour' vs 'color', or /ˈskɛdʒuːl/ vs 'schedule' for that matter; and there can be variations in articulation across speakers, speech communities, and even contexts, and standards change over time. Thus, it would be more proper to talk of standard articulation*s* at a given time and within a given community. However, I will continue talking of standard articulations, for simplicity. The thought is that not *every* attempted articulation of 'schedule' results in its utterance; e.g., articulating it as /bəˈnɑnə/ won't do. For more on the distinction between expressions and articulation, see Lepore (2009), Hawthorne and Lepore (2011).

[2]Kaplan (1990) thought intentions were *sufficient*. His position is slightly modified in Kaplan (2011) to respond to some of the worries concerning limits of tolerance. For criticism see Cappelen (1999); Hawthorne and Lepore (2011). The latter authors, however, tout tolerance type considerations.

[3]It is not impossible to both hold that intentions are sufficient and include limits of tolerance for deviation. E.g., one can say that an *appropriately manifest* intention is sufficient for word tokening, where an intention is *appropriately manifest* in a context, just in case it's recognizable in that context given the limits of tolerance. Such a move is quite familiar from the literature on intentionalism about context-sensitivity resolution (e.g., King 2014).

[4]The relativization to a local linguistic community is meant to allow that small variations tolerated within a community at a particular time can build over time or place leading to larger changes in articulation, as with changes in pronunciations across history, or across different dialects. Relativization to a context allows accommodation of specific circumstances, or knowledge about specific speakers (Hawthorne and Lepore 2011).

expresses this thought. To understand me, you have to recognize the words uttered, so that you can recover their meaning. And to respond, you have to token words that express your response, in this context, in a language we share. The entire endeavor crucially relies on recognizing utterances as being of particular words.

While linguists typically posit more fine-grained theoretical notions than a word, they still take words seriously: they take them to be distinguished by, e.g., their syntactic, semantic, and phonological features, and to be the domain of morphological and phonological processes, or the primary locus of stress and tone assignment (Bromberger 2011). Observations about utterances of words are in part what constitutes the empirical evidence that is the input to such studies; we rely on inferences from features of utterances of words to features of words they are utterances of (Bromberger 1989). And finally, both word individuation, and judgements about whether tokens are of the same word have been deemed relevant for a variety of philosophical debates; e.g., in discussions concerning puzzles about belief ascriptions (e.g., Kripke's (1979) Paderewski puzzle), theorizing about meaning-change (e.g., Evans' (1973) 'Madagascar' challenge to the causal theory of reference) or in theorizing about slur terms (Anderson and Lepore, 2013, *inter alia*). The topic is far from esoteric.

2 | UTTERANCES OF WORDS

2.1 | Intentions and repetition

To reuse a word you heard or used before, Kaplan maintains, it suffices that you *intend* to repeat the word. While "[we] depend heavily on resemblance between utterances and inscriptions… in order to divine these critical intentions […and] also take account of accent and idiolect and all the use clues to intention […i]t is the latter that decides the matter" (Kaplan, 1990, 104). One way to understand Kaplan is that the requisite intention is *to repeat* a particular performance of a word (Hawthorne and Lepore 2011). More precisely:

- **C**(onstitutive)**R**(ole of)**I**(ntentions): If someone intends to produce the word w used in a prior performance, then what comes out of their mouth (or from their pen) is a performance of w.

To ensure its utility in practice, CRI should be coupled with what Hawthorne and Lepore (2011) dub *Pervasiveness of Repetition*: for any non-neologizing performance p of word *w* there must be a prior performance of *w* the agent encountered that she intends to repeat in producing p. But as Hawthorne and Lepore contend, since CRI (together with Pervasiveness) requires speakers to remember a performance they intend to repeat, it is wildly implausible. A more plausible criterion drops the reliance on *repetition*.[5] Aside from neologizing, one utters a word if one intends to utter it:

- **(CRI*)**: If someone intends to produce the word *w,* then whatever comes out of their mouth (or from their pen) is a performance of *w*.[6]

[5]For reasons that will become apparent in §2.3, CRI* is not quite the condition Kaplan (2011) proposes. However, I shall argue Kaplan's amendment to CRI* is also unsatisfactory. Hawthorne and Lepore do not consider CRI*.

[6]There could be further complexities in characterizing the requisite intention, as an agent could have conflicting intentions. They might simultaneously intend to use the first word they heard today, and to use *that* word, where 'that word' denotes 'dog' in their mental lexicon, because they mistakenly believe that was the first word they heard today. Then one must say which intention is relevant.

As with CRI, since CRI* is a sufficient condition, its utility will depend on whether we normally have such intentions. Kaplan maintains we do, and one might find this obvious: after all, don't we normally intend to utter the words we use, and intend to use those that convey our message? Indeed, aren't intentions *necessary* for using a word? Even if my sneeze sounds like /wɜrd/, I have not thereby uttered the word 'word' or, indeed, any word. Similarly, even if the wind sounds like /wɜrd/, no word has been tokened.

The emerging picture is as follows: Kaplan maintains that if a speaker intends to use a word, say, 'affect,' then they tokened 'affect,' even if they misspelled or mispronounced it. These intentions are accessible to speakers and cede them authority over which words they token. Kaplan compares the question of whether a performance is a repetition of a word used or heard before, with the question of whether a song one performs is a variation of an existing song, or a new one, and writes "the natural answer is intentionalist: it is the [one] that the player has in mind and takes himself to be improvising on. I am not sure whether to be somewhat less intentionalist about songs and dances and games than about words, but I am still pretty intentionalist" (Kaplan 2011, 526-7).

Both the necessity and sufficiency of intentions has been challenged (Cappelen 1999; Hawthorne and Lepore 2011). But the challenges generally take the form: take two indistinguishable inscriptions or sounds (say 'word' or /wɜrd/) one produced with an intention to token a word, and the other purely accidentally, perhaps through no human activity at all. It is bizarre, challengers contend, to think that the former is a token of a word, and the latter isn't: after all, I could use a randomly generated inscription that entirely by accident matches a sentence of English to convey the message that sentence would normally express (Cappelen 1999). Thus, these critics conclude, intentions are not necessary. Regarding sufficiency, even if I intend to utter the word 'word,' if what comes out of my mouth is a mere grunt, I have failed (Cappelen, 1999; Hawthorne and Lepore, 2011). In short, the challenge to necessity suggests that if something looks like a token of a word, it must be one; the challenge to sufficiency tells us that if something doesn't (sufficiently) look like a token of a word, it cannot be one.[7]

I agree that intentions are neither necessary nor sufficient. However, I shall argue that the types of counterexamples employed to show this are misguided. The requirement that an utterance satisfies some standard of tolerance for deviation from the canonical articulation is likewise not necessary. Nor is a match—even a perfect match—with the canonical articulation sufficient. Both considerations of intentions, and of tolerance, are misguided. First, I turn to the problem with intentions. The authors that have challenged necessity and/or sufficiency of intentions have done so for the wrong reasons.

2.2 | On lexical selection: An empirical interlude

Kaplan oversimplifies the process of word production in assuming there is a unique point in this process where a speaker forms the intention to use a particular word; and, even if there were some such point, it is implausible that the relevant intention would be transparent to the speaker or decisive with respect to the identity of the word tokened. The leading accounts of speech production assume a process with multiple stages and levels of representation (see, e.g., Levelt (1989), Levelt et al. (1999), Dell (1986)). Roughly, production begins with pre-linguistic conceptualization, pre-verbally conceptualizing a message to be conveyed. At this stage, the agent formulates a thought, but has not yet

[7]One can endorse the challenge to sufficiency or necessity without endorsing the other. One can endorse the view that intentions are necessary, but not sufficient, for tokening a word; in addition one's performance has to satisfy the standards of tolerance for deviation from canonical articulation. One can also hold that intentions are sufficient, but only appropriately manifest ones (compare fn. 3).

activated and selected linguistic items that express it. The speaker then undergoes a process of lexical selection, whereby an item from their mental lexicon connected to the target concept is selected to express it. The mental lexicon houses lexical items (so called *lemmas*) which provide syntactic and semantic information needed to compose the syntactic form of the item and combine it in a sentence.[8] (For instance, the item 'dog' specifies its meaning, and that it is a count noun with a variable diacritic for number.) The target concept activates the lexical item it is connected to; however, once the item is activated the activation also spreads to related concepts and their associated lexical items. (For instance, DOG will activate the item 'dog,' but the activation also spreads to related items, e.g., 'cat.') The selected item is the most active one.[9] The selected item then provides input to morphological, phonological, and phonetic encodings, the steps in production which, in turn, produce a series of representations, from the input lexical item, to a combination of morphemes, to the phonological representation of the word, and finally, to a phonetic-gestural score, which is used in articulating the word.[10]

This process allows for errors at multiple junctures, and different types of speech errors arise depending on where the glitch occurs. For instance, semantic substitution errors can originate in conceptualization, if the speaker mistakenly focuses on a related but non-target concept in conceptualizing the intended message (e.g., CAT *vs* DOG); they can also arise in lexical selection, since conceptualization typically activates more than one of several related elements in the mental lexicon, and during the process of lexical selection the most active one is selected. Sometimes the most active one can deviate from the target, resulting in error: e.g., a speaker produces 'tennis bat' instead of 'tennis racquet,' or 'cat' instead of 'dog.' By contrast, sound-exchange errors happen during the process of phonological encoding, where the correct phonemes are produced, but some are wrongly positioned—e.g., 'big feet' vs 'fig beet.'

My point is that even if we describe lexical-selection as intentional, it is implausible to posit a unique intention that both determines which expression is tokened and is transparent to the speaker. Suppose a speaker makes a selection error, whereby the item 'bat' instead of 'racquet' got selected. In this case the speaker intended to token 'racquet' but ended up tokening 'bat.' One might think that the speaker must have also had a direct intention to token *that*, where 'that' on this occasion denotes 'bat.'[11] But, on standard accounts, substitutions occur when a competitor is sufficiently activated to

[8]Sometimes the notion of a "lemma" is reserved for syntactic information, the semantics being specified by the "lexical concept," each connected to its own lemma (see e.g, Roelofs, 1992; Levelt et al. 1999). The selection then begins with an activation of a lexical concept, which activates the corresponding lemma. The result of the initial stage of word production—lexical selection—still represents semantic and syntactic information, which is input to morphological, phonological, and phonetic encodings. The distinction doesn't affect our discussion.

[9]In other words, lemmas are connected in a network, and selected through a spreading-activation process: the most active one is selected. The result is input to further processing stages.

[10]While there are differences between models of speech production, what is uncontroversial is that production involves multiple stages, with multiple layers of representation. The differences, while important, are insignificant for our purposes, as my main point holds regardless. In brief, key differences between the main competitors are in whether they maintain modular, *top-down (serial) processing*, where the next stage of processing begins only after processing at a higher level ends, and information flows downstream (Levelt, (1989), Levelt et al (1999)), or they involve *parallel-processing*, non-modular, cascading model of production, where processing at lower levels can begin before the previous stage ends, so information can flow both downwards and upwards (Dell 1986).

[11]If they intended to select 'bat' because they mistakenly believe 'bat' applies to racquets, they both intended and selected the target 'bat,' in intending to talk about racquets. But it suffices to make trouble for CRI* that a non-intended, non-target item can be selected; and it can.

beat the target—*regardless of intention*. The item that outcompetes the target is precisely *not* intended. So, it's implausible to posit an intention to token 'bat;' the non-intended word was tokened.

And *even if* one were to posit an intention to token 'bat' in such case—an assumption independently implausible—it is unlikely such intention to token the non-target item would be transparent, nor is it plausible it would determine which item was tokened. It would co-exist with a number of other primary or "higher up" intentions—e.g., to convey something about racquets, to use a word denoting racquets, or even to use the word 'racquet.' We typically report an intention to convey "what we had in mind," so speakers would likely report they intended to use 'racquet,' or say something about racquets, but made a mistake. A "higher up" primary intention to token 'racquet' cannot suffice to secure that the speaker tokened 'racquet,' as this would require saying that even though the item 'bat,' but not 'racquet,' was selected and featured in the speaker's representation of the utterance, somehow, the speaker still tokened 'racquet.' This is implausible, and fails to capture that, despite the speaker's intentions, the targeted item was *not* selected—an error *was* made. Yet any putative direct intention to token 'bat' couldn't determine which item was selected either: the error was made in tokening the *unintended* item. Nor would it be accessible to the speaker, granting them authority in word individuation; we couldn't, with Kaplan, say the speaker uttered "what they had in mind" in any interesting sense.

In any case, regardless of what the speaker intended—even if they intended 'racquet'—if 'bat' was chosen during lexical selection, the speaker tokened 'bat,' just as in intending to pick up your racquet if you accidentally grab a bat, your action is one of grabbing a bat. Intending to token 'racquet' is no more a guarantee you will token it than intending to grab a racquet is a guarantee you will grab a racquet. What establishes that the speaker is tokening 'bat' rather than 'racquet' is that the former was selected; and that, we have seen, is not a matter of intention. Speakers might report they intended 'racquet;' but while such reports might serve as cues in interpretation, they do not determine which form was tokened. In short, one can fail to token 'racquet' even if one intended it, and one can token 'bat' even though one didn't intend to token it. Faced with these considerations, it is implausible that intentions are necessary or sufficient for successfully tokening an expression.[12] But, then, what is? I suggest this is a matter of lexical selection—the expression tokened is the one chosen during the selection process. This has nothing to do with intentions.

2.3 | Originalism-plus-Transfer and Tolerance

Here is a first pass proposal: which word an utterance tokens is a matter of which lexical item is selected in its production. Two utterances token the same word just in case the same item from the speaker's mental lexicon was selected in their production.[13,14] This, however, immediately raises two

[12]Downplaying the role of intentions does not render it easy to master a language; important differences separate, e.g., English speakers and parrots. Speakers use words, qua words of a language, to express a meaning. Parrots do not: they mimic sound. Further, they lack humans' cognitive capacities in language production (they lack a human mental lexicon and language faculty), nor can they be appropriately causally/historically situated, something which we will see below is of importance. Pace Millikan (1984), parrots or recording devices do not token words.

[13]We are considering only utterances that are products of a selection of a single item here.

[14]Recall, lexical items (lemmas) contain semantic and syntactic information that specify how a word can be combined with other expressions in a sentence. So, there will be a single lexical item for uninflected and inflected forms of a word (say, 'dog' and 'dogs'). The proposal thus yields a desirable outcome that an utterance of 'dog' (resulting from a selection of that item) is of the same word as an utterance of 'dogs ' (resulting from a selection of the same item).

worries. First, the condition might appear too strong. Since after selecting a lexical item plenty of errors can occur at downstream stages of production, it seems we allow utterances deviating significantly from standards of articulation to count as tokenings of a target word. Suppose one attempts to utter 'dog,' selects the correct item, yet due to a gross failure of articulation, produces a grunt. How could this count as an utterance of 'dog'? So, selecting an item from the mental lexicon might not be sufficient.[15] Mustn't an utterance also meet certain external standards of articulation to count as a token of a (particular) word?

This is the same type of worry that has been raised for sufficiency of intentions. It appears to motivate Tolerance (Hawthorne and Lepore, 2011):

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

Tolerance secures that a grunt is not an utterance of 'dog:' it does not meet the standards for performance within the (local) linguistic community, which places constraints on admissible performances of a word.[16] Such constraints are context-sensitive, and community specific. The locality condition is meant to permit larger-scale deviations from canonical articulations across (spatially/temporally) distant communities.

The second worry with my first pass proposal is that it tells us nothing about when distinct speakers utter the same word. We are told utterances are of the same word when they select the same item from the speaker's lexicon. But to answer when utterances by distinct speakers are of the same word, we have to say what makes two items in the speakers' respective mental lexicons correspond to the same word. I take up these worries in turn.

The problem with Tolerance is that it conflates metaphysics and epistemology. To see why, consider someone who writes "the noise effected my performance." The audience can hypothesize that, perhaps, the speaker misspelled 'affect(-ed),' as the words resemble each other phonologically and orthographically. Or maybe the speaker mistakenly uttered 'effected:' a slip of the tongue, or they mistook 'to effect' to mean *to affect*. That what was uttered deviated from a canonical articulation will not establish the speaker failed to have uttered anything (or accidentally uttered something else[17]). Inferring which type of error the speaker made is a part of the epistemic task of interpretation, a process of figuring out which form was tokened.

[15]Is it necessary? Recall the case where one finds an inscription matching a string of English words, say, "Nice day today," produced accidentally, through no human intervention (Cappelen, 1999). Couldn't I use it to make an utterance? I could succeed in conveying the message, but that doesn't mean I tokened any words. In presenting the note, I can speaker mean in Grice's sense that today is a nice day, even if the inscription contains no signs of English (Grice, 1957). Further, my intention will be appropriately manifest to my audience due to the accidental match between the inscription and the canonical articulation of the words.

[16]As noted earlier, Tolerance is widely endorsed to the point of being considered common-sensical. Even Kaplan, who ultimately rejects it in its full force, concedes that grunts cannot count as performances of words. This is the sense in which he deviates from CRI*: he concedes that performances so drastically unsuccessful as to result in mere grunts cannot count as performances of words, despite intentions. For endorsements of Tolerance see e.g., Cappelen (1999), Hawthorne and Lepore (2011), Sainsbury and Tye (2013, 59).

[17]It is unclear whether Hawthorne and Lepore (2011) take an intention to be necessary for uttering a word. But if both such intention and Tolerance were necessary, then, assuming 'affect' and 'effect' are sufficiently apart by local standards, the speaker who intends to utter 'affect' but misarticulates it as 'effect' would not have said anything. It appears to me (and Kaplan 2011) that this is an unfortunate consequence.

To object that 'effect' resembles 'affect' sufficiently to satisfy local standards for tolerance misses the point. Local standards can be useful guides for the *epistemic* task of interpretation but will not ensure you performed a particular word. The same holds for gross misarticulations. If in attempting to use 'dog,' one grunts, the audience faces the same interpretive task as with 'affect(-ed)'/'effect(-ed).' Maybe the speaker grunted, or was trying to draw attention to something, or they misarticulated a word. If the audience thinks it was a word, they can speculate which one was misarticulated. That the interpretive task might be harder than in the previous case does not render it of a different ilk. The challenge remains epistemic. Tolerance has no effect on the metaphysics of establishing a particular expression was tokened.[18] In short, Tolerance is misguided.[19]

One might worry that in rejecting Tolerance one is rendering communication mysterious: if gross misarticulations can be tokenings of a word, how can we reliably recognize the words uttered? While my account does allow for misarticulation, don't forget that our speech production mechanism is generally reliable, and that our utterances normally do not deviate dramatically from standard articulation in ways that makes the interpretive task impossible, or prohibitively hard.[20]

But if which word was tokened is just a matter of which lexical item is selected, how will we square this with intuitions about what is said by an utterance, when an error transpires? Consider Reverend Spooner, uttering "You have tasted the whole worm," instead of "You have wasted the whole term." Suppose, as likely, the correct items had been selected, but a sound-exchange error occurred. Did Spooner *say* that they wasted the whole term? Or, consider someone uttering "My grandmother lied" meaning 'died.' Suppose that in one case they selected the right item yet made a downstream error; in another they accidentally selected 'lied.' Do we predict that despite having uttered one and the same string, they said their grandmother died in the former case, and that she lied in the latter?

My response is not to confuse *saying* with tokening. *Says that* reports are notoriously promiscuous and context-sensitive, and often fail to track the literal meaning of an utterance (Cappelen and Lepore, 1997). In a context in which I want to convey why Spooner was angry at his students, I might correctly report him as having said that they wasted the whole term. In a different context, where I want to describe a funny incident, I can report that he said they tasted a whole worm. That both contexts are possible shows we shouldn't put too much weight on intuitions about *says that* reports. Similar considerations apply in the second example.

One can of course ask what the literal meaning was. But, normally, in interpreting speakers, we are trying to discern what they intended to convey. Take again "My grandmother lied." Suppose we realize an error was made. We can hypothesize that the speaker accidentally tokened 'lied,' in which

[18]The metaphysical question is not about what caused the speaker to token an expression, but what must be so for an utterance to token a particular expression (i.e., what determines whether the speaker tokened an expression x, misarticulated y, or failed to token anything.)

[19]Kaplan (2011) concedes too easily to Hawthorne and Lepore that grunts will not normally count as utterances. Of course, if someone is so incapacitated they are incapable of even *mis*articulating expressions, they will have failed to utter anything, inasmuch as they will have failed to perform an action altogether. Perhaps, they planned an utterance, but could not execute it, much like attempting to raise one's hand but suffering a sudden paralysis. They failed to raise their hand, even though they intended to. But if they are performing an action—uttering something—they might have dramatically bad articulations, and still be uttering something. Dramatic errors in articulation are still errors in articulation. Further, intuitions about the difference between not saying anything (failing to perform *any action*) and misarticulating dramatically, might themselves be driven by the epistemic considerations guiding interpretation, and may fail to reflect an answer to the question of which action was in fact performed, if any.

[20]If anything, the evidence suggests that the type of errors we make in word production are *not* random, and differ in kind depending on how they arise (see, e.g., Dell 1986; Levelt et al, 1998, i.a.). We know which types of error to expect, and by and large recognize them. That we don't witness mere grunts or random outputs all the time is no accident.

case the literal meaning of the form tokened would indeed be that the grandmother lied. However, that's irrelevant: after all, we will have understood that this was an accident and that the speaker meant *died*. We would have to engage in exactly the same task of trying to figure out what they meant if they misarticulated 'died' as /laɪd/, or if they produced something that was completely scrambled (e.g., producing a grunt instead of 'lied.') Regardless of the type of error—accidentally tokening something with an unintended literal meaning, or accidentally scrambling the articulation of something that has the intended literal meaning—the interpreter's task will be to figure out that the speaker made an error and that they meant that the grandmother has died.

Turning to the second worry with the proposal, we still need to say what makes utterances by two speakers of the same word. This requires us to say when items in their respective mental lexicons correspond to the same word. Here is a proposal. Words are coined via an act of neologizing. Neologizing might, but need not, be baptismal, as in a parent pointing to a child, saying "I am naming you…". The neologizing typically establishes the meaning of an expression.[21] The expression is stored in the neologizer's mental lexicon and ready for future use. Once stored, it can be transmitted through a chain of causal/historical connections to other agents, introducing it into their mental lexicons, thus creating a network of users.[22] Membership in this network allows speakers to use the word, and express its meaning.[23]

With this in place, we can maintain that utterances u and u' are of the same word just in case the items tokened in their production are appropriately connected, via a chain of causal and historical links to the same act of neologizing. More precisely, whether utterances token the same lexical item is a matter of whether the same item from the speaker's (or speakers') mental lexicon(s) was selected in their production; that is, whether the item(s) selected were introduced into the speaker's (or speakers') mental lexicon(s) by experiencing a token of the same word. And a token of which word a speaker experienced when they acquired a word will be a matter of which word was uttered by whomever introduced them to the word; this will again be a matter of the causal history of the item selected from the mental lexicon of the individual who produced the token that introduced the speaker to the word. In short, it is a matter of causal/historical connections. That is, I propose Originalism-Plus-Transfer (OPT):

- **O**(riginalism-)**P**(lus-)**T**(ransfer): Utterances u and u' by a speaker (or speakers) are of the same word just in case u and u' token (select) the same lexical item from the speaker's (or speakers') mental lexicon(s), grounded in the same network of causal-historical chains leading back to the neologizing event.[24]

[21] *Typically*, because I allow that sometimes meaning can fail to be established. I'm neutral as to *how* (or *if*) meanings get fixed, or what the nature of the meaning is.

[22] If which word is uttered is a matter of the selected item from the lexicon, can one token words one doesn't yet have stored in their lexicon? And if not, isn't this problematic, since we often appear to repeat words in the process of learning them? Luckily, we can address this worry. Mental lexicon is a part of the long-term memory. The process of lexical selection involves an interaction between the long-term memory and (verbal) working memory where the selected lexical item is recalled and processed. Repeating novel words, and word learning, involve a similar interaction between the long-term and working memory. So, it's plausible the representation one has in the working memory when attempting to store a word within the lexicon plays the role of the selected lexical item in production of familiar words. This process must involve a *specific* kind of interaction between the working memory and the long-term memory (engaging *verbal* working-memory). If one is just repeating a sound one heard, not processing it as linguistic, this wouldn't suffice for word learning. For similar reasons, parrots and recorders can't utter words: they don't have the appropriately set cognitive interface required for word learning and production to begin with. (Compare fn. 12 above.)

[23] This is akin to Kaplan's (1990; 2011) account, in that words are introduced and transmitted through chains of causal/historical connections between speakers. However, I am not committed to Kaplan's metaphysics of words, according to which words are concrete continuants of inter- and intrapersonal stages. If one allows that abstracta can be created, one can accept all I've said and maintain that words are abstracta.

[24] We are still considering only utterances that are products of tokenings of a single item.

3 | THE GENESIS MISTAKE

OPT appeals to an event of neologizing. But there are familiar problems with views that appeal to such notions. They have been raised as a challenge to a particular answer to the question of when *words* are the same, namely Originalism:[25]

- **Originalism:** w and w' are the same word just in case they are introduced by the same neologizing event.

Originalism naturally goes with OPT.[26] But a common challenge to Originalism is that it cannot make intuitive sense of changes in word histories (e.g., accounting for identity of words that share origin, or words that have multiple origins).[27] I call these *the problem of switch cases*. I believe it rests on a mistake about neologizing—one I dub *the Genesis Mistake*.

3.1 | Switch cases

Hawthorne and Lepore (2011, 483) invite us to consider a linguistic community, x, which uses the word 'happy.' Two travelers, from causally isolated communities c_1 and c_2, pass by x, and acquire this word; they return to their homelands and introduce their communities to it. Suppose the two communities come to articulate the word slightly differently, say, 'harpy' (/ˈhɑrpɪ/) and 'hapry' (/ˈhæpri/). It is natural, Hawthorne and Lepore maintain, to suppose c_1 and c_2 are using the same word slightly differently. But suppose further they also come to attach different meanings to 'harpy' and 'hapry.' Hawthorne and Lepore contend, one would have to say that, from the point of view of the members of x, c_1 and c_2 are using *the same word with different meanings*. Yet, they argue, from the standpoint of c_1 and c_2, they are using *distinct words*—a conclusion that might be forced on them: for if on some future encounter, c_1 users pick up 'hapry' from c_2, and c_2 'harpy' from c_1, members of both communities will maintain they are using different words, and "even if they learn that there is a common origin, this likely will not affect that judgment" (Hawthorne & Lepore, 2011, 438).

Let's take a member of c_1 and c_2, and let them utter 'harpy' (/ˈhɑrpɪ/) and 'hapry' (/ˈhæpri/), respectively.[28] Given OPT aren't we forced to say that they tokened the same word, since both tokens are connected in a single causal/historical network going back to the neologizing event?

Switch-cases are only puzzling if we presuppose what I call the *Genesis model* of word introduction via "canonical" baptisms, akin to Adam pointing at animals, saying "Call this thus-and-so," or "By definition, let 'x' denote the square root of 7." But the canonical case is not the only possible, and not even the dominant, way in which words can be introduced: they can be introduced tacitly, on the

[25]Among Originalists are Millikan (1984) and Richard (1990). Sainsbury and Tye (2013) defend Originalism for concepts.

[26]Given OPT one could say that w_1 and w_2 are the same word just in case any utterances of w_1 and w_2 must be of the same word. If all words are utterable, this would render a version of Originalism. There is a further question whether some words might be impossible to utter. And at the same time, we'll see, a word can have an origin though it needn't ever be uttered. (I return to this below.) In any case, I'm mainly concerned with establishing OPT. Thanks to Zoltán Gendler Szabó for discussion.

[27]As we'll see shortly, Hawthorne and Lepore (2011) raise this type of worry for Originalism, and for word-realism more generally.

[28]Assume that their utterances didn't involve any misarticulations, and they selected the targeted items from their respective lexicons.

fly, in conversation, even without the interlocutors realizing the word is novel. Consider neologizing through zero-derivation, where a new expression is created from an existing one, preserving the phonological shape, but not the lexical category of the original, as when one says of a mysteriously missing prisoner: "He houdinied his way out."[29] Such examples are ordinary. They rely on agents coordinating on a new expression-meaning pairing and establishing a novel convention through such coordination.[30] Agents can also co-opt a previously existing *articulation* of an expression in introducing a new expression-meaning pairing. That is, a group can come to coordinate to introduce a novel expression with a novel meaning by co-opting an existing articulation. In the case above, the articulation of the name 'Houdini' is exploited in zero-derivation. An existing articulation can also be coopted if you get named after someone or something else: e.g., being named 'John' after John Locke, or 'River' after rivers. In such cases, an articulation of an existing expression is co-opted in establishing a novel name—your name.

With this in mind, we can easily explain the 'harpy' vs 'hapry' case. At some point, 'happy' got introduced into a language, fixing its meaning. It is then transferred through the linguistic community by a network of causal/historical connections; its meaning transferred as well (Kripke 1980). That c_1 and c_2 come to articulate 'happy' in slightly different ways does not yet have an effect on word identity. But once the communities settled on a new meaning convention, they each neologized a word, coopting old articulations in establishing novel expression-meaning pairings, and so, from these neologizings on, we have two words, 'harpy' and 'hapry,' with different meanings. The mistake was to implicitly adopt Genesis: neologizing can occur tacitly, in communication, and it can exploit previously existing expressions' articulations. The community might start using the existing articulation (a string 'harpy,' pronounced (/ˈhɑrpɪ/)) to coordinate on a particular meaning, either knowingly or not. When communities c_1 and c_2 started using articulations 'harpy' (/ˈhɑrpɪ/) and 'hapry' (/ˈhæpri/) to coordinate on different meanings, new conventions were established, and two new words (canonically articulated as 'harpy' (/ˈhɑrpɪ/) and 'hapry' (/ˈhæpri/)) were created. If so, then, in our example, the members of c_1 and c_2 aren't tokening the same word, after all. They are selecting lexical items that trace back to *different* neologizings.[31]

[29]Adapted from Clark and Clark (1979). See Armstrong (2016), Davidson (1984, 1986, 1991) and Lepore and Stone (2017) for further discussion of lexical innovation, and for references on this issue. Zero-derivation is just one of many possible instances of lexical innovation.

[30]See Armstrong (2016) and Lepore and Stone (2017), on the role of coordination in lexical innovation "on the fly." The idea of coordination as a prerequisite for convention goes back to Lewis (1969). Note, conventions can start in smaller linguistic subcommunities, and spread to larger ones—we can coin a word long before it spreads to become a "standard" word of English. More on this below.

[31]I claimed a change in meaning conventions leads to a neologizing of a new word. But we often speak of words changing meanings, and this change is often perceived to be gradual. Insofar a replacement of one word by another is not gradual, it might appear my view clashes with the intuition. But those who claim that words change meanings are no better off: insofar as change is gradual, there would be a period of time during which it is vague, or there's no fact of the matter, whether the word has the new or the old meaning. This doesn't square well with intuitions; if anything, while the new use has arisen but the old hasn't yet died out, we typically observe a pattern of ambiguity, suggesting multiple lexical items in play. (Consider the "new" usage of 'nauseous:' "Both I and the smell of my medicine were nauseous," is at best a pun, suggesting there's an ambiguity between the "old" meaning, and the "new" one. 'Nauseous' doesn't behave as polysemous, nor do we feel there's no fact of the matter what it means. Similarly with, e.g., 'cool,' meaning awesome, which developed from 'cool' meaning cold.) On a positive note, my account *can* accommodate the intuition of a gradual change. Conventions often start locally, among small subgroups of a linguistic community. Local conventions can develop rapidly given appropriate future-directed commitments (c.f., Armstrong 2016; Lepore and Stone 2015) and then gradually spread throughout the wider community. During gradual spread, there might be a period when two words exist (accounting for ambiguity intuitions), the original one eventually falling out of use. (More on words dying out below.) Thanks to Zoltán Gendler Szabó for discussion.

The theoretical advantages of this explanation become transparent once we look at actual cases of historical change. English 'parent' and Italian 'parente' both trace back to the Latin 'parēns'. While 'parent' means parent, 'parente' means relative. If pressed to say 'parent' and 'parente' are the same word (with different articulations) due to their same origin, we would have to say a single word has both meanings simultaneously. If this were so, we should be able in general to access both on either articulation—something we cannot do. Utterances of "All relatives are parents" or "My parents are also parents of my parents," have no readings which exploit putative multiple meanings of 'parent.' It is a part of the lore that where there are multiple unrelated meanings, not simultaneously accessible, as with ambiguity (e.g., 'bank'), there are multiple lexical items at play, and not one word with two meanings. Insofar as cognates are like this, we should say that, through the establishment of new conventions, multiple words are created from the articulation of an existing one.[32]

Further, if 'parent' and 'parente' were a single word with different meanings in different linguistic communities, it would be difficult to explain what happens when these communities introduce one another to their respective usages. It would be strange if it suddenly became possible to truthfully say "All relatives are parents," as a result of this encounter. Nor, as already emphasized, is it plausible 'parente' would suddenly pattern as a single item with multiple meanings, even upon such encounter. "Both my mother and all my relatives are my parents," has no true reading that exploits the putative different meanings for 'parent,' regardless of whether I speak both Italian and English, though it should if there were a single lexical item with multiple meanings at play. One might think one can only access the respective meanings via a specific articulation. But we know this is problematic: articulation is inessential to the identity of a word being tokened, so how can it be essential to the *meaning* being expressed?

Still, does it not offend intuition that a new meaning being associated with an existing string creates a new word? In the 'harpy'/'hapry' case, isn't there also an intuition that members of c_1 and c_2 are using the old word 'happy' with a changed meaning—an intuition that might come naturally to the members of the original community x? After all, we do talk as if words change meanings:

1. 'Lunch' used to mean one thing—any meal—but now it means the mid-day meal.
2. 'Parent(e)' means something different in English and Italian.

Shouldn't we say the *meaning* of 'lunch' changed by establishing a new convention, rather than that a novel *word* 'lunch' got created once a new meaning was conventionalized?

[32]My account thus yields an attractive solution to putative reference-shifting problems for the causal theory of reference. Evans (1973) famously argues 'Madagascar' changed its reference, due to an error by Marco Polo, from Mogadishu to the African island. On my account, once the community started to coordinate with Polo in 'Madagascar' referring to Madagascar, a new convention was established, coopting an old articulation (Polo's misarticulation of the original' name for Mogadishu) to coin a new expression with a novel meaning. As with 'harpy'/'hapry' case, this happened implicitly, without Polo or anyone else intending to coin a new term. The establishment of the new convention was a neologizing to which our common day tokenings of 'Madagascar' trace down. This is why the name we token is of the island. Since contemporary speakers were typically not exposed to the old name, they lack a name 'Madagascar' for the mainland in their mental lexicons. So, they lack access to an expression articulated as 'Madagascar' (/ˌmædəˈgæskər/) denoting Mogadishu. I explore this approach to putative meaning-shifting puzzles in more detail elsewhere.

(1)–(2) strike us as plausible precisely because they track articulations.[33] The problem with holding that these judgments track words, instead of articulations, is that they extend to *false* cognates. Just as we say 'parent' means different things in English and Italian, we can also say 'brat' means different things in English and Serbian, even though they share no common origin. The same goes for false cognates within one language: we say 'bat' means both a club and a flying mammal of the order *Chiroptera*, but one disambiguation has its roots in North Germanic, and the other in Old English, and likely Celtic. Similarly, we can report a word has *the same* meaning in two languages even for false cognates: in English and Mbabaram, 'dog' means dog. But this is pure coincidence. With false cognates, there is no temptation to posit a single item—it is a pure coincidence the words share an articulation. So, intuitions do not discriminate between sameness of articulation and sameness of words. As with 'parent' and 'parente,' the intuition that in the 'harpy'/'hapry' case the original community *x* might describe c_1 and c_2 as using their word 'happy' slightly differently treads on articulations, bolstered by the common ancestry.

As per OPT, then, we shouldn't conclude that Italian and English speakers uttering 'parente' and 'parent,' respectively, (while selecting the corresponding items from their respective lexicons), are tokening the same word. Nor should we conclude that the speaker today, and the one of the past, are uttering the same word, when they utter 'lunch.' Our word 'lunch' was coined by coopting the articulation of the old word 'lunch.'[34] But the two trace back to different neologizing events. Intuitions to the contrary track *articulations,* not words.[35]

But need there be a unique neologizing event? Could a single word be introduced at two distinct times? If causally isolated communities introduce what superficially appears as one word, we are dealing with multiple words (recall Mbabaram vs English 'dog'). If they came in contact, even though

[33]This might also underscore the intuitions about reports such as "my name is the same as yours." Hawthorne and Lepore (2011) argue against Kaplan's *common currency* names (as opposed to *generic names*, which can apply to many individuals, e.g., 'John' as a name for anyone thus named), by pointing out that names do not behave as ambiguous expressions, and that reports of sameness of names track intuitions about generic names. But there are reasons to think names *are* ambiguous, just as they would be on a common currency account. While it is true that, as they point out, "I am Sam and he is too" is felicitous, "I met Sam and I greeted him" (where 'him' *is* anaphoric), or "I met and greeted Sam," have no reading where I met and greeted two distinct individuals named 'Sam,' just as expected if 'Sam' were ambiguous. I lack space to explore this issue in detail.

[34]If there is a new item 'lunch' after the change, shouldn't we be able to access the older one even after the change? Not necessarily. It is entirely possible for words to become obsolete, and gradually less frequently tokened, even extinct (think of extinct languages). Since to acquire a word one needs to be exposed to it, it is only to be expected that through time fewer individuals are exposed to the original word, until it is no longer a part of the mental lexicon of (most) present day speakers. (Compare fn. 32.) And while it is, as we have already seen, we tend to witness the pattern of ambiguity. (Recall fn. 31.)

[35]I argued that establishing new meaning conventions involves neologizing words. This explains why false cognates, or etymologically related, yet nonsynonymous, words are distinct. But what about *synonymous* cognates? English 'telephone' and Italian 'telefono' are cognates *and* synonyms. Are they the same word? And are their utterances of the same word? There are grammatical—syntactic—features that distinguish them: e.g., only the latter has a grammatical gender marker. Recall, lexical items (lemmas) in the mental lexicon carry semantic and syntactic information. It is natural to say that when there is a change in not just semantic, but also syntactic conventions, a new word is coined. But we should resist that changes along *just any* grammatical features result in neologizing. Changes in phonological features don't: one can speak different dialects of a language; one can acquire an accent, or loose or change one, often tacitly. If all those changes added new words to the mental lexicon, this would quickly lead to an explosion. Allowing for such variations to automatically result in neologizing repeats the mistake of shape theorists (e.g., Davidson, 1979; Cappelen 1999) who, in identifying words with their articulations—their phonological or orthographic form—fail to capture that a single word can be (mis-)articulated in a myriad ways. (A similar worry applies to "feature-bundle views," e.g. Chomsky (2000).) Since lemmas store only syntactic and semantic information, the line I draw between semantic and syntactic, and other kinds of changes, is principled. Thanks to Zoltán Gendler Szabó for pressing this point, and thanks to Michael Glanzberg for a related discussion.

the words are articulatorily indistinguishable, still two words remain (it is a coincidence they came in contact). That they might be able to understand one another's tokenings of 'dog' is no objection to OPT or Originalism. They might have good evidence that they are using the same word 'dog,' and since by accident the words are synonymous, they can recover the intended meaning. But that they are interpretively lucky in this way is no objection to OPT or Originalism.

What if different members of *one* community attempted to independently coin 'dog' to mean *dog*? Plausibly, some uptake from a community is required for convention, so only once the convention is established within the community is the word successfully coined. If so, the neologizing will coincide with establishing of a convention, not with the first performance of a string sounding like /dɔg/. In short, switch cases are not a problem for OPT, nor for Originalism either![36]

4 | CONCLUSION

Contrary to common assumption neither intentions nor Tolerance are either necessary or sufficient for determining which word an utterance tokens. Instead, I defended what I dubbed Originalism-plus-Transfer account. By separating metaphysical and epistemic considerations concerning the role of speaker intentions and Tolerance, my account captures the intuitions that underscored the appeal to both, while offering a more satisfactory answer to what determines which word an utterance tokens. Further, I have shown that OPT should be coupled with a more nuanced understanding of neologizing—one that departs from Genesis. This understanding of neologizing also allowed us to dispel the common objections to Originalism about word individuation. Thus I argued, a promising account of word individuation is available as well.

ORCID

Una Stojnić https://orcid.org/0000-0003-2151-2462

BIBLIOGRAPHY

Anderson, L., & Lepore, E. (2013). Slurring words. *Noûs*, *47*, 25–48. https://doi.org/10.1111/j.1468-0068.2010.00820.x

[36]It has been argued that a further worry for Originalism is posed by *unperformed* words: if words are introduced by neologizing, then there cannot be unperformed words (Hawthorne and Lepore, 2011). Since OPT is only concerned with *utterances* of words, this is not automatically problematic for OPT. But even for Originalism, the problem is misguided. That there is no first performance, doesn't mean that the word wasn't coined: neologizing needn't be a performance. I can introduce a word by description, without uttering it. Of course, nothing guarantees I will succeed. If in 2021 I utter: "Let the word consisting of the concatenation of the names of two firstborns in 2022 in NYC mean…" and the world explodes, I'll have failed to introduce a word; but likewise if I said, "I name you 'Adam'," unless I had the authority to name the addressee. Similarly, it's no problem for Originalism or OPT that there are words derived from others through derivational morphology, e.g., adding a prefix ('un-', 'il-') to an adjective. Languages vary in how systematic their morphology is—not any combination of an adjective with a prefix is a word. Where systematicity is lacking, introducing the combination explicitly is needed. And insofar as morphology is highly systematic, the multiplicity of new words—those derivable from combining with relevant morphemes according to the derivational rule—is introduced with the introduction of such rule. Nor need Originalist say all derivates of a single rule are the same word. One could treat the introduction of a derivation rule as multiple (simultaneous) events of neologizing, or instead say that words derivable by systematic morphological rules are individuated relative to simple words they combine *and* the derivational rule(s) they invoke. In the latter case, one can accommodate this within OPT by saying that two utterances are of the same word just in case they either token the same simple item from the respective speakers' mental lexicons leading to the same neologizing event, or their production involves selecting simple base items leading to the same respective acts of neologizing, to which the same morphological rules have been applied in the same way.

Armstrong, J. (2016). The problem of lexical innovation. *Linguistics and Philosophy*, *39*, 87–118. https://doi.org/10.1007/s10988-015-9185-9

Bromberger, S. (2011). What are words? Comments on Kaplan (1990), on Hawthorne and Lepore, and on the issue. *Journal of Philosophy*, *108*, 486–503. https://doi.org/10.5840/2011108925

Bromberger, S. (1989). Types and tokens in linguistics. In A. George (Ed.), *Reflections on Chomsky* (pp. 58–90). Oxford: Basil Blackwell.

Cappelen, H. (1999). Intentions in words. *Noûs*, *33*, 92–102. https://doi.org/10.1111/0029-4624.00143

Cappelen, H., & Lepore, E. (1997). On the alleged connection between indirect speech and the theory of meaning. *Mind and Language*, *12*, 278–296. https://doi.org/10.1111/j.1468-0017.1997.tb00075.x

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: CUP.

Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, *55*, 767–811. https://doi.org/10.2307/412745

Davidson, D. (1979). Quotation. *Theory and Decision*, *11*, 27–40. https://doi.org/10.1007/BF00126690

Davidson, D. (1984). Convention and communication. *Synthese*, *59*, 3–17. https://doi.org/10.1007/bf00873278

Davidson, D. (1986). A nice derangement of epitaphs. In E. Lepore (Ed.), *Truth and interpretation: Perspectives on the philosophy of Donald Davidson* (pp. 433–446). Cambridge: Blackwell.

Davidson, D. (1991). James Joyce and Humpty Dumpty. *Midwest Studies in Philosophy*, *16*, 1–12. https://doi.org/10.1111/j.1475-4975.1991.tb00228.x

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*, 377–388. https://doi.org/10.2307/2182440

Hawthorne, J., & Lepore, E. (2011). On words. *Journal of Philosophy*, *108*, 447–485. https://doi.org/10.5840/2011108924

Kaplan, D. (1990). Words. *Proceedings of the Aristotelian Society*, *64*, 93–119. https://doi.org/10.1093/aristoteliansupp/64.1.93

Kaplan, D. (2011). Words on words. *Journal of Philosophy*, *108*, 504–529. https://doi.org/10.5840/2011108926

King, J. C. (2014). Speaker intentions in context. *Noûs*, *48*, 219–237. https://doi.org/10.1111/j.1468-0068.2012.00857.x

Kripke, S. A. (1980). *Naming and necessity*. Cambridge, MA: HUP.

Kripke, S. A. (1979). A puzzle about beliefs. In A. Margalit (Ed.), *Meaning and use* (pp. 239–283). Dordrecht: Reidel.

Lepore, E. (2009). The heresy of paraphrase: When the medium really is the message. *Midwest Studies in Philosophy*, *33*, 177–197. https://doi.org/10.1111/j.1475-4975.2009.00190.x

Lepore, E., & Stone, M. (2017). Convention before communication. *Philosophical Perspectives*, *31*, 245–265. https://doi.org/10.1111/phpe.12103

Lepore, E., & Stone, M. (2015). *Imagination and convention*. Oxford: OUP.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38. https://doi.org/10.1017/s0140525x99001776

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lewis, D. K. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT Press.

Richard, M. (1990). *Propositional attitudes: An essay on thoughts and how we ascribe them*. Cambridge Studies in Philosophy, Cambridge: CUP.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*, 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Sainsbury, M., & Tye, M. (2013). *Seven puzzles of thought: And how to solve them*. Oxford: OUP.