In "Words" (Marantz (2001), and elsewhere in the same time period) I claimed that some derivational affixes could attach either to roots or to categorized roots, i.e., to roots plus an affix that provides a syntactic category. Doublets such as "donor," from the root √DON, and "donator," from the root √DON + verbalizing "ate," supported this idea. However, it was never clear in what sense the two uses of such affixes constituted two uses of the same affix, as opposed to the use of two different affixes that shared phonological form. In particular, how can the syntax of these uses be the same when one affix attaches to a phase (a categorized and, therefore, interpreted stem) and the other does not? If roots themselves don't have meanings, how does the combination of an affix with a root have a similar compositional semantics to the combination of the "same" affix with a categorized root, which does have (a range of) meaning(s)?

Here I will describe a theory of morpheme meanings that will help us address this question of the possible identity between and among derivational affixes. I will not provide an answer to the question, are there one or two o/er derivational suffixes, but I will provide the framework in which we can give that question some empirical force.

Let's begin with some basic assumptions. I will adopt the general structure of Distributed Morphology; in particular, I will assume that the syntax combines abstract morphemes which map onto phonological representations at the PF interface and onto semantic representations at the LF interface. Agreeing (in general terms) with Borer against some versions of DM, I will assume that the underlying phonological form of a root morpheme serves as its identifying and individuating label in the syntax (see a forthcoming Blog post specifically on this notion). As far as the syntax is concerned, this label, although made up of phonological features, could have been a number, as in Harley's conception of roots, since the phonological features are not referenced by any syntactic rule or principle. However, since the grammar does map these labels onto semantic representations, speakers can identify statistical correlations between sound representations and meaning representations that might influence their assignment of meanings to new roots. (See the extensive recent work by de Zubicaray (2024, and the references cited therein) documenting these correlations and speakers' tacit knowledge of them.) Although I'm siding with Borer against Harley on the identification of roots via their phonological form, I'm also in agreement with Harley contra Borer that there are real examples of root suppletion across languages. The implication of the phonology as label for roots assumption for root suppletion is that suppletion is directional, replacing the underlying form with a marked form in a particular environment. Thus I'm in partial agreement with Nanosyntax, which treats root suppletion as overwriting of a "smaller" lexical entry by a "larger one" (the larger entry includes the context for the suppletion), although I reject the superset principle and phrasal spell-out. That is, root suppletion does involve "overwriting" such that the phonological form of a root might involve a set of forms. I'll return below to the connection among rule-based contextual allomorphy, morphologically specific phonological rules/constraints, and root suppletion below (or in a further post).

I will also adopt the contextual embeddings/vector-based theory of meaning being explored in computational linguistics, including work on Large Language Models. Here, a meaning of a unit in sentential context is a vector (an ordered list of values) specifying the position of the unit's meaning in a multidimensional meaning space whose dimensions are related to the contexts in which the unit has appeared in in a (super large) corpus. Each token of a unit in a corpus has its own vector. To the extent that the contextual meanings of multiple tokens share meaning, they will cluster in this meaning space, occupying spaces near each other and apart from the spaces that other units map to.

Adopting this view of meaning is more than a matter of convenience for me, although for most of what I want to do in this paper, the details of the contextual embeddings framework won't matter. Quine, in *Word and Object* and other work, explained why we should think that meanings of words are learned via their distribution in sentences, rather than from a correspondence between language and the perception of the outside world. His "meaning holism" is perhaps the best explanation of the ability of Large Language Models to generate meaningful sounding texts. However, it's important to remember that the dimensions of embedding space correlate well with the dimensions of meaning space derived from "embodied" approaches to meaning (see, e.g., Fernandino et al. 2016 – this approach has been more fully explored for Cognitive Neuroscience by Binder's group). That is, the brain is structured around its input and output systems, and the processing of input/output relations seems to structure the space that learning through distribution "discovers." There is no basic incompatibility between the notion of a genetically determined embodied semantic organization, an empiricist view that meaning is learned via (some) interaction with the environment, and a meaning holistic/distributionalist view that the meaning of linguistic units is overwhelmingly a distillation of their distribution in language. Everyone's right about meaning, but the devil is in the details.

The idea of meaning vectors is not really different from the old school feature-based meaning theories of, e.g., Fodor & Katz in the generative tradition. A set of features like, [+/-human], [+/-male] is a meaning vector describing a point in multidimensional feature space, a fact that appealed to structuralists like Jakobson (the structure of the Russian case system is a cube! (using semantically grounded binary case features)). The main special properties of the semantic features of a standard generative grammar are that they're binary, rather than valued between 0 and 1, and they're likely innate, rather than learned via an analysis of a large corpus. Much of the points below would be valid under the assumption of binary innate features as opposed to continuous valued learned features.

The convenience of modern contextual embeddings as model of the meaning of linguist units is that the models have various implementations that can produce numbers for experimental investigations. That is, I can quantify "similarity in meaning" for empirical work. In addition, there has been work mapping contextual embedding space onto neural architecture, allowing for linking hypotheses between Marr's algorithmic and implementational levels of analysis within Cognitive Science and for explicit hypotheses about where to expect what in the brain during MEG experiments.

In the mapping from syntax to LF, then, a unit (leaving vague which unit for the moment) is mapped onto a point, or a set of points, in a multidimensional space whose dimensions are generalizations over shared environments of units in the language. A basic assumption of DM is that roots require Merger to a category node, e.g., a v, n, a or p (for verb, noun, adjective or preposition – in other work I have argued for a decomposition of these category nodes into a minimal set of features, but the standard "little x" categories will suffice for the present discussion). Work in DM on meaning in context argues that the meaning of a root is "fixed" in the context of the first category node with which it combines. I will take that proposal to mean that roots themselves don't have meanings; only the combination of a root plus a category affix has a meaning (note the resonance here with Aronoff's lexeme-based morphology assumptions). We can make that claim concrete within the contextual embeddings approach to meaning: roots don't map onto embedding space at LF by themselves (the null context isn't a context for contextual embeddings for roots) – a root needs another unit for a contextual embedding.

Using our example of "rotor, rotate," we see that if roots don't mean anything outside of the context of a category head, then the grammar itself does not make any demands on a shared meaning between "rotor" and "rotate" – the n affix (-or) and the v affix (-ate) each map the root onto a place in meaning space, but nothing requires that these places are connected. This should be the same situation faced by pairs like "(to) clash, (the) clash," noun/verb pairs formed by zero little n and little v affixes attaching to the same root. If the noun/verb pairs here are related in meaning, we say that the root is polysemous and that the two words are "senses" of a single unit. If a noun/verb pair is not thought to involve a meaning relationship, as in "(to) duck, (the) duck," the root is said to be ambiguous and the two words are homophones and have different "meanings" (as opposed to "senses").

The literature on polysemy details a distinction between systematic polysemy, e.g., the relation in English between the food and animal meanings of words like "duck" and "chicken," and non-systematic polysemy, like, perhaps, the meanings of words like "calm" across emotions and weather. Let's call systematic polysemy "r-polysemy" for "rule-based polysemy," and assume that speakers know/learn the relevant rules. These rules describe a mapping between locations in meaning space, e.g., between the space where animal meanings are clustered and the space where food meanings are clustered. When Al Capp introduced the animal, the "shmoo," into his comics, he could write about eating some shmoo and readers would be able to create the r-polyseme meaning for the noun using their rule.

We take the semantic connection between "rotor/rotate" and "clash/clash" to be one of r-polysemy. As such, a speaker encountering one member of the pair in context need not recover the meaning of the other member of the pair as part of the semantic analysis of the sentence being processed, as with eating chicken or shmoo.

In “rotor,” there is a contribution to the meaning of the noun from the suffix, -or. A “rotor” is a thing that does something – an instrument, like (some) other -or words. We expect to find the meaning of “rotor” in the meaning space for instruments like “carburetor” and “transistor,” as well as -er/or nouns made on verbal stems, like, perhaps, “governor” and “farmer.” How would we represent the meaning of derivational affixes that attach to categorized roots, like “farm” as a verb? Since a categorized root will map to a location in meaning space, we would want such a derivational suffix to map from the verb meaning space of its stems to the -er meaning space of the derived nouns. That is, there should be a “rule” or generalization relating these meaning spaces that is similar to a r-polysemy rule. Let’s call this a derivational meaning rule, or DMR for short. There are several differences between a DMR and an r-polysemy rule. First, the DMR is directional, from the stem meaning to the meaning of the derived form, whereas the r-polysemy rule need not be directional. Second, a DMR is necessarily part of the derivation of a derived form built on a categorized stem whereas a root plus derivational suffix combination may not be related to another word via an r-polysemy rule (just like a food name need not be related to an animal name via r-polysemy, e.g., “beef”). Third, the meaning of a derived word based on a categorized stem is compositional, which means that the meaning of the stem is still part of the complex representation of the derived word, whereas word related via r-polysemy in no way contain the meanings of the words to which they are r-polysemy-related.

At this point we can define a way in which two derivational suffixes, one that attaches to roots and one that attaches to categorized stems, can be called the “same suffix” in a way that might make predictions (that is, we could test whether the two are analyzed as one within a grammar). First, they would have the same root “label” (same phonological form, as a root). This characteristic depends on analyzing at least some derivational suffixes as being combinations of a root and a category node, as I have argued for recently (along with others). Second, they would map their stems onto the same region of meaning space (the “donor” words co-located with the “farmer” words in meaning space). Third, the r-polysemy rule relating the root+suffix word (donor) to a root+suffix word (donate) with the same root would need to resemble the DMR that connects the stem of the categorized stem + suffix word to the stem (so the r-polysemy between donate and donor would resemble the DMR that takes $farm_v$ to farmer). The factor that would drive a learner to ask the question, are these two things the same, is the phonological identity between the two candidate suffixes. Given identical roots, do the other two properties hold? Given a root + category combination, like √ER/OR + n, we ask, does this combination map on the same region of semantic space, independent of the semantics (or lack of semantics) of what it attaches to? If yes, what’s the DMR that connects the semantics of the stem, when it has a semantics, to the semantics of the derived form? Then, could that DMR be turned into an r-polysemy rule connecting the root-based derived forms with the suffix to words of the appropriate category (verb, in our example) built from the same root? For any given root base, the lack of a root-based derived form of the appropriate “source” category isn’t a strike against considering the root-taking suffix and the stem-taking suffix as the same suffix – the lack of a verb from the stem “butcher” doesn’t count against considering the -er in “butcher” to be the same suffix as the -er in “farmer,” since “butcher” should be in the same semantic region as “farmer” and the other occupation -er words.

[What about the -er in “father,” “mother,” “sister,” and “brother”?  One could argue that the consistency of the semantic space occupied by this group of relative words could be the basis of an analysis in which the root -ER + n suffix maps roots onto semantic space.  Then is this the same -ER as in “donor” and “farmer”?  Here see the discussion of polysemy vs. homophony in a following post:  yes, they should be the same suffix, meaning the same root and same category head, with “sister” and “donor” having the same syntactic analysis. However, the -ER + n suffix could map to two distinct portions of meaning space, meaning that the two “usages” of -ER + n are homophonous rather than polysemes.  Whether there is an interesting meaning relation between the occupational -ER and the familial -ER such that they are in fact polysemes (is there an r-polysemy rule to state here?) is another question.]

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., ... & Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral cortex*, *26*(5), 2018-2034.

Marantz, A.  (2001).  Words.  Widely distributed handout from WCCFL 20, USC.

de Zubicaray, G. I., Kearney, E., Guenther, F., McMahon, K. L., & Arciuli, J. (2024). Statistical relationships between surface form and sensory meanings of English words influence lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*.