Words, Species, and Kinds

JTM Miller

Forthcoming in *Metaphysics*; Pre-proof Version – please cite published version

Abstract: It has been widely argued that words are analogous to species such that words, like species, are natural kinds. In this paper, I consider the metaphysics of word-kinds. After arguing against an essentialist approach, I argue that word-kinds are homeostatic property clusters, in line with the dominant approach to other biological and psychological kinds.

Pygmy three-toed sloths, also known as the monk sloth or dwarf sloth, are between 48 and 53 centimetres long, and fully grown adults weigh between 2.5 and 3.5kg.[1] These are true statements concerning the species or kind, pygmy three-toed sloth. Here is a simple argument for the existence of the species, pygmy three-toed sloth: assuming the above claims are true, they quantify over (or at least refer to) species; we should accept the existence of entities that are quantified over by true statements; therefore, the species, pygmy three-toed sloth, exists.

A parallel case can be made for the existence of word-kinds. Words are similarly quantified over within true claims, such as ''Paris' contains 5 letters'. This claim is not about any token word, but about the word-type or kind 'PARIS'.[2] Such 'type-talk' is ubiquitous in both in our scientific and ordinary talk about words, and in both we constantly refer to (and quantify over) words qua kinds. As we want to hold that these claims are true, this provides a prima facie case that we should accept the existence of words qua kinds.

Indeed, an appeal to a direct analogy between words and species is very common within the metaphysics of words literature.[3] It appears in its most developed form in Wetzel (2009):

> 'Words should be viewed as real kinds, just as species are' (2009: 5)

> 'If we accept those sentences as true, then we must acknowledge the existence of generic entities such as species, genes, epigenotypes, languages, syllables, vowels, allophones, quartets, chess openings, atoms,

---

[1] For more details on these fascinating and beautiful animals, see Anderson and Handley Jr. (2001). This is also recognized as the first description of the pygmy three-toed sloth.

[2] In this paper, the terms 'type' and 'kind' can be taken to be interchangeable. Most literature on words uses the notion 'type', but I will favour 'kind' to make the comparison with literature on natural kinds easier.

[3] See Miller (2020b) for an overview. My focus is on essentialism within this ontological or metaphysical approach to words.

and so on, all of which were apparently copiously quantified over' (2009: 26)

> 'Nonetheless, I will argue that just as members of a species form a kind, so do the tokens of a word. It will be my contention that the word is what glues its tokens together, that words are important nodes in linguistic taxonomy just as species are in zoological taxonomy' (2009: 106).

As we will see below, those who have endorsed this analogy have very different opinions about how to individuate these kinds (or types), but most have accepted that words are like species and that word-kinds exist. However, securing the existence of word-types leaves open questions about the nature of such types or kinds. The dominant view in the metaphysics of words literature has been to adopt a form of essentialism about words, but I will argue that this faces serious problems. Instead, I propose that words are best thought of as homeostatic property clusters.

1. The Metaphysics of Species

Essentialism about species, drawing on a tradition back to Aristotle, holds that species are natural kinds, defined by their real essences, statable in terms of necessary and sufficient conditions for kind-membership. All instances of a species (or kind) have a common essence, in that they all instantiate some property (or properties) that ensure they fulfil the necessary and sufficient conditions of being an instance of that kind. There are many well-known issues with essentialism about species, some of which I discuss more fully with reference to essentialism about word-kinds below.

The homeostatic property cluster view (HPC), first developed by Boyd (1990, 1991, 1999a, 1999b, 2000, 2003a, 2003b), aims to reject the idea of providing necessary and sufficient conditions on kind membership, whilst maintaining that groups of entities share some cluster of projectable properties. The central motivations for the HPC view as originally stated by

Boyd are epistemological. The HPC view seeks to explain how the kinds that we refer to 'in inferences, explanations, and predictions are groups of entities or phenomena that are similar to one another but not perfectly the same' (Ereshefsky and Reydon 2015: 970). Hence it attempts to explain the typical traits of a kind, whilst rejecting the necessary and sufficient conditions required by essentialism.

Under the HPC view, kinds apply to entities that instantiate properties where those properties are contingently clustered in nature. This clustering of properties occurs as, 'at least typically,

the result of what may be metaphorically (sometimes literally) described as some sort of homeostasis' (Boyd 1999b: 143) such that the presence of some properties tends to favour the presence of other properties within the cluster, or certain underlying mechanisms maintain the presence of properties within the cluster. This clustering of properties explains the causal powers of kinds. As members of a kind instantiate properties from the same cluster, those members will be causally relevant in similar ways.

The mechanisms that underlie the property clusters are important for the HPC view as the view's supporters reject the idea that the cluster of properties *alone* can provide a definition of a kind. This is because of the well-observed variation in properties among members of a kind that mean that the relevant property cluster can only be identified after the kind has already been relatively well established (Ereshefsky and Reydon 2015: 971). Therefore, under the HPC view, it is the combination of the cluster of properties, *and* the mechanisms underlying those properties that defines a kind and explains the epistemic role that kinds play within our theorising about the world.

It will be important to note, as Ereshefsky and Reydon do, that 'a homeostatic mechanism can be anything that causes (in the broadest sense of the term) a repeated clustering of properties' (2015: 971). This means that the relevant mechanisms should not be a priori restricted to just those that appear in certain sciences or are investigated in certain ways. As the definition of kinds is a posteriori (Boyd 2000: 54), we should also remain open to changing our view on what mechanisms, or what properties, define some particular kind. The kind may be able to preserve its identity despite a change in either the property cluster associated with that kind, or the mechanism that underlie that property cluster (Boyd 1999b: 144).

Applying this view to species, under the HPC view, species cannot be defined through necessary and sufficient conditions. This is because we cannot find properties of traits that are uniquely shared by all and only members of a single species. Instead, under the HPC view, a species is defined by a cluster of properties and the underlying homeostatic mechanisms that cause and sustain those properties to cluster. Members of a species thus share many, but not necessarily all, properties that are caused by these mechanisms, which may be grounded in various genetic and/or environmental mechanisms. Species under this view are therefore real. Species (and other kinds) track genuine causally relevant distinctions in the world, though which species (or kinds) we take to exist may to some extent reflect our theoretical perspectives, and the species (or kinds) that we posit may change as our theories change over time.

My aim in this paper is to extend this view to the case of words. I will argue that the HPC view provides a natural way to understand word qua kinds, accommodating our intuitions about how words can change, and how the instances of a word can vary in properties whilst still being an

instance of the same word-kind. In short, the HPC view will allow us to maintain a realist account of word-kinds, whilst accepting that those kinds are messy, historically delimited, non-eternal, and allow for exceptions – all features that I will argue word-kinds intuitively possess.

2. An Important Distinction

Before discussing words directly, we need to be clear on two different ways we might talk about words and kinds. The first is the kind WORD. This kind includes as its instances all written, signed, and spoken utterances. [4] The second is some particular word-kind such as 'STOCKPORT' or 'TABLE', whose instances are particular written, signed, and spoken utterances such as the written tokens: 'Stockport' and 'table'. To try to make this distinction clear throughout, I will adopt the convention of using capitals to indicate kinds (e.g., TIGER; WORD), and lower case to indicate particular entities (e.g. [a] tiger). When applied to words,

I will use quote marks. Thus, 'TABLE' refers to the particular word-kind, and 'table' to some particular instance or utterance of that word-kind.

Intuitively, we tend to think that members of the kind WORD are ontologically distinct in some important way from members of the kinds WATER and TIGER. This is, presumably, why we do not routinely mistake instances of the kinds WATER and TIGER for being instances of the kind WORD, and why we have such strong intuitions against bizarre naturally-formed rock formations being words despite any resemblance between the rock formations physical properties and the properties of some word.

This distinction has consequences for how we understand 'nominalism' about words. One form of nominalism about words might deny both that there are particular word-kinds and that there is some metaphysically genuine distinction between instances of words and instances of water and tigers, thereby rejecting the existence of the kind WORD. However, it is at least logically

---

[4] There is, in fact, significant ambiguity in how we talk about the kind WORD. Sometimes, as suggested in text, we use the kind WORD to refer to all instances or tokens of words. Thus, when Andy Dufresne in *The Shawshank Redemption* says, 'Get busy living, or get busy dying', we think that there are seven instances (or tokens) of the type WORD as there are seven particulars uttered. However, we might alternatively take the kind WORD as being the kind that has as its instances the more specific word-kinds. This is akin to the way that the dictionary contains words as it contains a list of word-kinds (such as the word-kinds 'STOCKPORT' and 'TABLE'). Or, we could take the kind WORD to be the kind that has its instances other still other kinds like the kinds NOUN and ADVERB. Some combination of these is also possible. Nothing in the arguments in this paper turns on this distinction as it is not my aim to argue what should be the proper extension of the kind WORD. My thanks to the anonymous reviewers for this journal for pushing me to discuss this in more depth.

possible for us to posit one of these kinds but not the other. We might deny the existence of particular word-kinds while accepting the existence of the kind WORD.

To see this, consider a view that holds that word-types like 'STOCKPORT' and 'TABLE' do not exist. This form of nominalism is only locally nominalist if accepts that there is some kind of entity, WORD, that is distinct from other kinds of entity such as WATER and TIGER. Perhaps the theory would be that all words share some property that makes them instances of the kind WORD but reject that all instances of some particular word share any metaphysically important similarities. The converse is also possible. We might accept that words are kinds, whilst denying that word is a kind. Under this view, we might think there is no metaphysically significant kind WORD – i.e., nothing that makes all words instances of the kind WORD – whilst accepting that there are particular word-kinds like 'TABLE'. On this view, then, all instances of 'table' are instances of the kind 'TABLE', but the kind WORD would presumably be (at best) a useful way of talking about the world.

This is not to defend these forms of nominalism – indeed, it is not immediately obvious what benefits we might get from some of the logically possible views. Rather, my intention is to clarify this distinction to allow us to clearly see what the main topic of this paper is. My focus is on the more specific kinds like 'TABLE' and 'STOCKPORT', and unless otherwise specified I will use the term word to refer to such kinds, not to their tokens. Nothing I say here will directly talk to whether the kind WORD exists. Perhaps that kind does exist and can be accounted for in a similar way as I will propose we can account for 'TABLE' and 'STOCKPORT', but I leave a direct discussion of that possibility to later work.

3. Why might there be word-kinds at all?

Before I discuss the specifics of metaphysical approaches to word-kinds, it is important to consider why such entities are posited at all – what phenomena they are taken to explain. As noted above, a first reason why we might posit word-kinds is that we quantify over word-kinds in a similar way to how we quantify over species. Just as there are true statements about the properties of species where no individual of that species need have all of the relevant properties, similarly there seem to be true statements about the properties of words (qua kinds) where no instance of that word need have all of the relevant properties. No pygmy three-toed sloth need be between 48 and 53 centimetres long, or weigh between 2.5 and 3.5kg, but the statements about the species are still true. It is true to say that ''Paris' contains 5 letters'. In both cases, the claims are true because they are about kinds, not particulars (as there may be instances of the word-kind 'PARIS' that do not contain 5 letters, such as in cases of misspelling).

Another common argument in favour of word-kinds is that they allow us to accommodate variation across instances or individuals. Just as individuals may vary in their properties whilst still being members of the same species, instances of the word 'colour' might vary in their semantic, orthographic, grammatical, and phonetic properties.[5] Indeed, it might be that no instances of the word share all the same properties while still being, intuitively, instances of the same word. It is often claimed that positing word-kinds allows us to explain this, as we need not think that every instance of a kind must share all the same properties whilst still being instances of the same kind. (See, e.g., Wetzel 2009.) As we will note below, whether word-kinds succeed in accommodating this variation is a debated point, but for now it suffices that a common aim in positing word-kinds is to be able to accommodate variation of instances of the same word.

Another reason to posit word-kinds is that they provide a case against the possibility of what I will call a 'grueified language'. From Goodman (1955), we can say that the property of being grue is not law-like and is not projectible. It is this that forms the basis of his new riddle of induction. Analogously, we can say that a grueified language would be a language where there is no explicit or projectable patterns and regularities within the language and the ways that its speakers use words within that language. This would mean that such a language would be one where a speaker of the language cannot be certain that a word has not radically shifted its properties, and that even two speakers who are otherwise fluent in the language in which they are conversing would not be able to be certain that the other person would be able to adequately interpret and understand each other's utterances.

Our natural languages are clearly not grueified in this way. Our linguistic practices rely on there being strong patterns and regularities at least in the sense that if I am in a shop and order a coffee, I assume that the instance of the word 'coffee' that I utter will be interpreted in line with relatively stable conventional practices and norms. Without these, it is hard to see how successful linguistic communication could proceed.

This stability also ensures that speakers cannot unilaterally radically change the uses of commonly accepted words. Speakers may of course attempt to introduce new words or change the accepted patterns of use associated with some particular word, but it is rare for some speaker

---

[5] Orthographic and phonetic variation might be easy to imagine, but how might instances have different grammatical or semantic properties? Consider an instance of 'colour' that is a noun and one that is a verb (as in 'to colour'). Such instances vary in their grammatical properties. On sematic properties, we might think that the context of utterance can alter the semantic properties of two instances of the word 'COLOUR'. Such cases will depend on how we define semantics, and how we understand the relationship between semantics and pragmatics. Thanks to an anonymous reviewer for raising this issue.

to be able to do this alone. Indeed, in cases where it may seem possible for a single speaker to invent a new word or to radically change the pattern of use of an established word, such efforts will only be successful if the new usage is adopted by the wider community of speakers. For example, we might think that Donald Trump was able to single-handedly invent the word 'covfefe'. However, the word has not been widely accepted, and many would not accept it as a legitimate use of language (in part as it is not clear what it was intended to mean, or if it was intended to mean anything at all). For the new word to be adopted into the language, the pattern of usage must be recognized by the community of speakers as sufficiently stable and established.

Of course, this is not to say that such patterns are absolute or unchanging. We must not posit kinds that are too strong in that they do not allow flexibility in language use. However, the patterns and regularities must be sufficiently stable so that speakers can reliably infer the intended meaning of words and sentences from the utterances of other speakers, and that their utterances can be also suitable interpreted. Positing word-kinds is often taken to provide this stability across time and speakers within a given community.

These are common reasons provided in favour of positing word-kinds. Often it is the case that none is alone intended to be conclusive, but together they are normally taken to be persuasive. It remains to provide an account of the nature of word-kinds such that they have the metaphysical characteristics to be able to provide the above benefits.

4. Essentialism About Words

Given the discussion above, essentialism may initially seem attractive. An essentialist view of words posits certain necessary and/or sufficient conditions for word-kindhood, such that all instances of some word-kind share certain features or properties. If essentialism is correct, then there is no threat of a grueified language as essentialism provides a strong enough account of words to make them projectible. There are, though, two main problems with the view that I will highlight here. The first is well-known in the literature and arises from attempts to specify what the necessary and/or sufficient conditions are for any given word-kind; the second argues that essentialism produces a conception of word-kinds that is too strong, and hence is unable to accommodate the messy historical nature of words.

Much of the recent literature on the metaphysics of words has focused on specifying the necessary and sufficient conditions on kind membership to provide the criterion of identity for words. For example, Kaplan (1990, 2011) argues that speaker intention is central to kind

membership such that tokens are tokens of the same word if and only if the speaker intends them to be.[6] Irmak (2018) argues that it is historical properties, detailed (at least in part) by the etymologies of words, which secures kind membership such that two tokens are tokens of the same type if and only if they have the same history. Hawthorne and Lepore (2011: 31) attribute to Mark Richard and Ruth Millikan the view that words are identical if and only if they have the same originating event. In each of these cases, what is being proposed is a criterion for kind (or type) individuation. In each case, particular properties of the token words are highlighted as significant for answering when it is the case that tokens are tokens of the same type. For each of these authors, despite their differences, a core aspect of the essentialist approach has been assumed: all the views try to specify certain necessary and/or sufficient conditions on word-kindhood. For all of them, words are like species, and both are defined through necessary and/or sufficient conditions on an individual being an instance of a given kind.[7]

All these proposals face serious problems. Cappelen (1999) has argued that intentions cannot be part of the individuation of words, Hawthorne and Lepore (2011) argue against individuating through originating events on the grounds that some words do not have a single originating event, and Miller (2020a) has argued against the use of historical properties as necessary and sufficient conditions. In each of these cases, the problem is that whatever we might try to argue are the necessary and sufficient conditions on word-kind membership, the variability of instances of the same word leads to problems. If there is no property that all tokens of a kind share, there is no plausible simple essentialist criteria for word-kind membership.[8]

Even some that recognize these problems in providing the necessary and sufficient conditions have maintained a commitment to essentialism about words. For example, Hawthorne and Lepore defend what they call 'sloppy realism' wherein 'unsettled questions [about word individuation] 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' (2011: 36). Wetzel accepts that

---

[6] More precisely, Kaplan's view is that particularly stages are stages of the same continuant if they have the same intentional (and some other historical) properties. Stages might vary in certain properties, but for two stages to be stages of the same continuant requires the utterer to intend to produce a new stage of that four-dimensional object. Kaplan also takes his own view to be naturalistic and opposes essentialism. However, Kaplan takes essentialism to be a form of Platonism. This is not the notion of essentialism I have in mind in this paper.

[7] Notable non-essentialist alternatives are Szabo (1999) and Nefdt (2019) though it is unclear what the metaphysical status of kinds are on their views.

[8] I say 'simple' as the view that I defend below, a homeostatic property cluster view, parallels views labelled as 'new essentialism' in the philosophy of biology. See Rieppel (2010) for a discussion of 'new essentialism' in the philosophy of biology, and Ereshefsky (2010) for some critical remarks.

'there is nothing interesting all and only uttered tokens of a particular word have in common other than being tokens of the word' (2009: 106-7). For Wetzel, kinds (or types) are abstract, eternal and unchanging, and the necessary and sufficient condition for being a member of a kind is to simply be an instance of the kind. It is clear from this that Wetzel maintains a commitment to the essentialist idea of necessary and sufficient conditions, wherein the necessary and sufficient conditions for Wetzel are that all tokens of a type have the property of 'being a token of type x'. Tokens may vary in relation to any and all of the other properties that they have: they are tokens of the same type so long as they possess this type-membership property. Hence, Wetzel is positing a brute instantiation relation between instances and kinds.

Sloppy realism, or an appeal to brute instantiation relations, though, is unlikely to be persuasive to those not antecedently committed to an essentialist account of word-kinds. The view lacks explanatory power in that it provides no way in which we could investigate and answer particular cases where we are initially unclear whether some instance is an instance of a given word. Is the instance 'apris', for example, an instance of the kind 'APRIL' or the kind 'PARIS'? Sloppy realism and brute instantiation relations provide no help in answering this question. Both seem initially plausible in cases where are intuitions are clear – of course 'color' and 'colour' are both instances of the kind 'COLOUR'. But, in more borderline cases, the view rather seems to embrace (at least epistemic) vagueness. The prospects for essentialism are therefore not promising. There is no plausible set of necessary and sufficient conditions, and the alternative to adopt some form of 'sloppy realism' is explanatorily unsatisfactory.

Even if we can provide a clear set of necessary and sufficient conditions, essentialism about words fails to strike the right balance between stability and flexibility in language. On one side, it makes word-kinds too strong and strangely ahistorical. If essentialism is correct, then word-kinds are defined by certain necessary and/or sufficient conditions. However, those conditions cannot change, else the kind would also change. Say that the kind 'TABLE' is defined by certain necessary and/or sufficient conditions, T. If, though, over time, the word 'TABLE' came to be used in new ways, the new instances may not satisfy those conditions. 'TABLE' may come to be defined by some distinct conditions, T*.

This might not initially seem like a problem. The essentialist might simply say that these are in fact distinct kinds - 'TABLE' and 'TABLE*'. However, there are many cases where words have undergone very significant changes over time and yet ordinary speakers (and linguists) still consider them the same words. Words are reappropriated, the use of them shift and change radically, and (as seen above) there is no good candidate set of necessary and/or sufficient conditions that can accommodate this within the essentialist view. Words precisely are changing, evolving entities, reflecting societal structures and the community of speakers that

use those words, contra the overly strong kinds posited by the essentialist. Words under the essentialist view therefore cannot provide the flexibility discussed above.

The supporter of essentialism may respond that word *tokens* reflect the interests and communicative aims of speakers. But this seems to misunderstand the importance of sociological influences on language and the ways words change and come to mean something very different, or to be spelt or pronounced in a new way. Appealing to a change in the tokens of the word, and not the word itself is (again) not in line with the ways that linguists or ordinary speakers talk about words and language. Ordinary speakers do not think that word change is purely a reflection on tokens of that word, but rather that the change in use of tokens of the word change the nature of the word itself.

For example, consider the word-kind 'EGREGIOUS'. In its original usage, tokens of this word meant something illustrious, but modern tokens mean something particularly bad. I suggest that ordinary speakers do not think that this change is solely explained by the tokens being used in a different way, but rather that the tokens being used in this different way leads to a change in the kind also, and it is this change in the kind that then dictates further usage patterns. Furthermore, if change is at the level of tokens, and not at the level of kinds, then kinds in fact play no role in the stability of language over time and within a community (see Miller 2020a). This goes against one of the initial motivations for positing kinds discussed in section 3.

Summarising, word-kinds were intended to provide stability needed to avoid the threat of a grueified language. The necessary and sufficient conditions posited within an essentialist view can provide that stability, but in so doing the view makes words ahistorical and rules out the flexibility we observe in language. Words are flexible and change, and this cannot be accommodated within the essentialist view.

5. Words as HPCs

From section 3, we can see that what we need are word-kinds which are strong enough to explain the projectible nature of words to avoid the threat of a grueified language. But, in section 4 we saw that we need kinds that are weak enough to accommodate the messy, historically delimited nature of words. I argue that the HPC view is well placed to balance these needs.

Applying the view to words, the position holds that word-kinds are defined by a cluster of properties, and the mechanisms underlying those properties. The immediate question is which properties. To see what I have in mind, consider our attempt to define the word 'TABLE'. When we look at accepted instances of this word, we see common clusters of properties, such as the

semantic property of expressing the semantic content <table>, the phonetic property of being pronounced /ˈteɪb(ə)l/, the orthographic property of being spelt 't-a-b-l-e'. If we think that words are entities that exist across languages, then the cluster may also include the properties of being spelt 't-a-v-o-l-o' and being pronounced 'ˈtavolo'. There may also be other properties within the cluster, in part depending on other positions we hold with regard to the properties of words more generally, but instances of 'TABLE' will (tend to) have a number of these properties.

Note also that this cluster may overlap with, but be distinct from, clusters that we find instantiated by instances of other words. Instances of the words 'DOG' and 'CAT' may show up in many similar linguistic contexts, but by considering a large number of instances of each word and the properties that each instance possesses, we can begin to notice differences in the clusters of properties associated with instances of distinct word-kinds.

However, under the HPC view, this cluster of properties is not enough to fully define 'TABLE', as there could be instances that have these properties but are not instances of the kind. Borrowing a common idea, it could be that an ant walks a pattern in the sand that is orthographically identical to instances of the kind, and yet we want to hold that the ant has not produced an instance of the kind. The mechanisms underlying the clustering of properties accounts for this. Given this, the kind 'TABLE' is defined through the cluster of properties associated with the kind *and* the mechanisms that cause that repeated clustering of properties. The immediate question will be which properties and what mechanisms?

Taking the latter question first, I suggest at least two different sorts of mechanisms are relevant in the case of words. The precise details of these mechanisms will depend on various empirical and theoretical commitments beyond the scope of this initial paper outlining the HPC view of words. For example, if you are an inferentialist (see Brandom 1994, Peregrin 2008) or an externalist (see Putnam 1975, Burge 1979), then you may not want to posit what I will call 'internal mechanisms' and may instead think external mechanisms are determinant of word-kinds. However, given my aim of providing a broad account of the benefits of an HPC view of words, pending future research, I will leave aside these more complex issues and provide an indicative sketch of what these distinct sorts of mechanisms might look like.

The first are 'internal' (or 'individual') mechanisms. These are mechanisms within a person's mind that are responsible for deriving sentences such that the lexical item 'table' is used in a regular and consistent way. Such mechanisms are internal to an individual and ensure that an individual can retrieve lexical items regularly and reliably from their mental lexicon. These mechanisms explain why a particular speaker does not routinely misuse the word in accordance with their own understanding of its meaning and patterns of use. These internal mechanisms are

also responsible for recognition effects such that a person can suitably recognize and interpret instances of 'table'.

This does not rely on any linguistic theory about how those internal linguistic systems (such as the mental lexicon) are structured or the elements that make them up. There are various ongoing debates about the precise nature of the mental lexicon, and the syntactic processes that are involved in the production of language. Nothing in this paper requires taking a specific view on those debates. What is required here is the widely accepted idea that all linguistically capable people have (broadly) similar mechanisms within their minds responsible for language. This is a minimal claim, and one that could only be rejected by arguing that different people have widely differing cognitive structures. The precise details about how sentences are derived within the mind/brain would be a matter for empirical studies in linguistics and cognitive science; all that is required here is that the process of derivation wherein lexical items are retrieved from the mental lexicon and combined syntactically is minimally consistent across derivations for a particular person, providing one part of the explanation as to why we observe the repeated (and relatively stable) clustering of properties.

Also, nothing here limits the mechanisms to human cognitive systems. Some animals can be taught sign language and hence be taught words within those languages. Plausibly, machines too can learn and then express words. Neural nets used in machine learning might be extremely different from human processing, and other animals may use and token words with varying cognitive structures. The HPC view of words, as it is defended here, need not hold that it is the *same* internal mechanisms in humans, animals, and machines. All that is needed is that the internal mechanisms, in each of these cases, ensure that lexical items are used in a regular and consistent ways.

The second sort of mechanism are 'external' (or social) mechanisms. These are the mechanisms that are responsible for speakers within a community following the accepted linguistic norms within that community. These mechanisms account for the accepted regularities across speakers and their use of a word. Such mechanisms will be, to a significant degree, socially constructed, and be responsive to the wider patterns of use within a community of speakers.

A good place to start to understand these external mechanisms would be to appeal to the tolerance principle suggested by Hawthorne and Lepore which states that:

> Tolerance: Performance p is of a word w only if p meets relevant local performance standards on w (2011: 18).

The external mechanisms that cause the clustering of properties can then be cashed out in terms of these 'local performance standards'. The precise nature of such standards would seem again to be a matter for empirical studies beyond the scope of this paper, but we do have some intuitive

grasp on some such standards from our ordinary language use. For example, there is a limit of tolerance for how to spell or pronounce words within a community of speakers in line with that community's performance standards.

The performance standards that are relevant to whether the words are being tokened by a person, animal, or machine, will not only be language-specific (e.g., specific to English or French): they will also be particular to specific communities of speakers (even, potentially, a community of speaker that includes a mixture of persons, animals, and machines). For example, the relevant performance standards maintaining the external mechanisms will vary in formal and informal contexts. Speakers need to follow these norms, and the norms are providing a mechanism that is in part external to the individual that ensures a certain level of regularity within the range of local performance standards tolerated by the community of speakers.

The combined effects of these mechanisms explain the clustering of properties, and hence account for the observed homeostasis of the property clusters that partly define word-kinds. Neither is sufficient alone for this, but together they ensure that each instance of the kind 'TABLE' will instantiate some relatively stable set of properties, such that, for example, utterances with different pronunciations are (within the limits of the tolerance principle) correctly classified as being instances of the same kind, and that speakers are able to create new instances of the kind regularly and consistently via the internal mechanisms underlying our linguistic abilities.

Returning to the first question, we can also ask which sorts of properties cluster. I have already mentioned orthographic, phonetic, semantic, grammatical, and perhaps intentional properties, but this list is not intended to be exhaustive or exclusive. Nor will all words have some properties of all of these sorts. Words in languages that lack any written form will lack orthographic properties entirely. We might think that at least semantic properties are more important than other properties. But even this will depend on further views on whether all words have semantic properties. For example, nonsense words might plausibly have no semantics, and yet instances of these kinds (assuming they are words at all) will still exhibit a clustering of properties. This is compatible with the HPC view which only holds that there are properties that cluster. I am not here taking a position on which properties are part of any cluster, or whether any are necessary or more important than other properties.

The HPC view is also compatible with the properties in the cluster being interdependent in other complex ways. For example, there may be a dependency between the semantic properties and the orthographic properties in some languages that is missing in others. The semantic and orthographic properties instantiated by instances of the English word 'EYE' cluster but there is no further dependency between then. Conversely, there is clustering and a dependency between

13

the semantic properties and the orthographic properties of the hieroglyph ' ⌒ '. Thus, the kinds

of properties that cluster homeostatically, and any dependencies between those properties, will be different for different languages, and even potentially within a language. The HPC view can explain this as these dependencies will arise due to the same mechanisms that creates the clustering of properties in the first place.

Adopting the HPC view of words has numerous advantages. First, it provides a principled reason as to why our natural language is not grueified. Words have projectible properties because the stable mechanisms cause (in the broadest sense of the term) repeated clusters of properties. For example, these mechanisms explain why the presence of the semantic property of expressing the semantic content <table> favours the phonetic property of being pronounced /ˈteɪb(ə)l/ and/or the orthographic property of being spelt 't-a-b-l-e'. This means that once we know what cluster of properties partly defines a word, and incorporated that knowledge into our linguistic system, we can safely project that future instances of a word will have the same properties when the new instances arise out of the same homeostatic mechanisms.

It also allows us to explain why the ant walking in the sand does not instance a word. They do not because, presumably, the mechanisms underlying the ant's movements are not the same as those that underlie our derivation of sentences and other linguistic structures. This account therefore allows us a way to deny that the ant's movement patterns instantiate a word token without an appeal to intention. This is not to say that the mechanisms or properties cannot include intentional aspects. Rather, it is merely a consequence of the HPC view that it allows non-intentionalist accounts a new way to respond to the kind of intentionalism defended by Kaplan (1990, 2011).

Second, this view allows that different instances may vary whilst being instances of the same kind. As kinds are defined by a *cluster* of properties, it need not be the case that all instances share the same properties. Whilst the view allows for instances of 'table' that vary to be instances of the kind 'TABLE', it also allows for extensional vagueness and fuzzy boundaries for the kind. The level of vagueness and number of boundary cases will vary depending on how established and stable the mechanisms are. For very new words where there are no clear settled mechanisms, this explains our split intuitions about whether these are cases of words. For example, take 'COVFEFE'. People are split over whether this is a word, and the HPC view can explain this as being due to the lack of established (internal and external) homeostatic mechanisms associated with the word, and hence a lack of a pattern of clustered properties. In most cases, like that of 'TABLE', though, the mechanisms are stable and well-established, and this explains why we are less likely to think that 'TABLE' is a fuzzy kind. But the view still

leaves open the possibility that even the most stable mechanisms and property clusters may be extensionally vague.

Third, as the homeostatic mechanisms are not individuated extensionally and are instead conceived of as historical entities that can change over time (Boyd 1990: 374, 1999a: 88, 1999b: 144, 2000: 71), words can change over time. Over time the mechanisms can shift, explaining why it is the case that the same word can come to be partly defined by a different cluster of properties. Words can evolve under the HPC view as the mechanisms themselves can cause different properties to cluster over time.

This also accounts for the observed historicity of words. Words can have different properties at different times because the homeostatic mechanisms that cause the clustering of properties are themselves relative to a time and may change over time. Mechanisms are relative to times in the sense that both internal and external mechanisms are themselves instantiated by entities (people and communities of speakers) that are located at a particular time. This suggests that the property clusters they cause are also best understood relative to a time.

This is not to say that the mechanisms *must* change. Perhaps some do not, and certainly some change faster than others. It is a well-recognized empirical phenomenon that certain words are more likely to change their properties, and this, I argue, is due to the mechanisms that underlie the clustering of properties associated with those words are less stable and hence more likely to change.

Finally, under the HPC view, whilst word-kinds are real, they are created. Words are created by the bringing into existence of new mechanisms. New mechanisms would result in new clusters of properties, and thus new words would be created. As above, how stable that new word is, or how widely accepted as a word it will be, will depend on the stability and (perhaps implicit) acceptance of the underlying mechanisms. Take 'COVFEFE' for example. Is this a word? Under the HPC view, the answer to this question will depend on whether there are underlying mechanisms that cause and sustain relevant property clusters. Pending further empirical investigation into the nature of the relevant internal and external mechanisms, I would suggest that in this case there are no such stable mechanisms. There is no clear set of linguistic norms for speakers within a community that ensure that instances of 'covfefe' are used in a consistent way, nor the relevant internal mechanisms ensuring that a speaker can instance the word in a stable consistent way. Given this, I would (somewhat cautiously) say that there is no such word-kind 'COVFEFE'.

However, were such mechanisms to come into existence that would result in the stable, projectible use of the word, then there could be such a word brought into existence.

Shakespeare, therefore, on this account did successfully create new words as his actions did result in the existence of stable mechanisms that cause and sustain property clusters to this day.

All these features, I argue, show that the HPC view of words allows us to maintain most, if not all, of the intuitive claims we have about words. They are created, but messy; historical but able to change. The projectibility of words is explained, and the threat of a grueified language avoided. Perhaps most importantly such kinds are real, tracking genuine causal effects in the world, and are studiable by the sciences through a consideration of both internal and external mechanisms. A greater understanding of those mechanisms would provide the potential to solve boundary cases where it is unclear what word is being instanced. And the HPC view allows words to be, in a sense, both private and public such that, once internal mechanisms are established, they are sufficient for the creation of certain clusters of properties but are also responsive to the external, social mechanisms.

6. The Ontology of Words and the Metaphysics of Language

What does adopting the HPC view mean for the ontology of words and the metaphysics of language more broadly? The first, likely unsurprising, consequence of this view is that we cannot accept it and Platonism about words together, or at least any sort of Platonism that embraces elements of essentialism that the HPC view seeks to reject. For (some) Platonists (e.g., Wetzel (2009); Katz (2000)), words are eternal entities, whilst for the supporter of the HPC view, they are created. The Platonist seeks necessary and sufficient conditions that are non-theory relative, whilst the supporter of the HPC view rejects this and in so doing attempts to explain the variability of words in a new way.

However, the HPC view better aligns with other positions defended recently in the literature, especially nominalist views. For example, take Bromberger's view that words are 'archetypes' (1989). Bromberger's view of words is intended to allow us to maintain type-talk without accepting the existence of abstract entities. Tokens are members of quasi-natural kinds, and types are archetypes (or models) of those kinds. As Wetzel puts it, types are 'defined as something that models all the tokens of a kind with respect to projectable questions but not something that admits of answers to individuating questions. Thus, for Bromberger the type is not the kind itself, but models all the tokens of the kind' (2006). As types are models which are 'object[s] so designed that, by finding the answers to some questions about it, a competent user can figure out the answer to questions about something else' (Bromberger 1989: 62), the view is nominalist.

This view can, prima facie, be easily combined with the HPC view, especially as 'no pair of objects stands (or fails to stand) in the model/modelled relation absolutely, but only relative to specific sets of questions, pairings of questions, and algorithms' (Bromberger 1989: 63). Models change relative to the questions and interests, in a similar way that we might think that the mechanisms that cause and sustain property clusters might change in virtue of our interests over time.

Alternatively, Miller (2021) has argued that words are bundles of properties. Miller argues that 'if tokens are bundles of properties, then types are bundles (or sets or collections or pluralities) of tokens, where those types have their criterion of identity in virtue of the properties of the tokens that are members of type' (2021: 5741). This ontology, if correct, would cohere very neatly with the HPC view. Indeed, the HPC view would additionally allow a bundle theory of words to explain *why* it is the case that particular bundles of properties are observed to regularly co-occur.

This is not to say, though, that the HPC cannot be combined with a more realist account, or that it settles the debate between the realist and the nominalist. Irmak (2018), for example, defends the idea that word-kinds are artefacts. This is consistent with the HPC view. Indeed, Boyd argues for the HPC view by saying that kinds are 'features of human inferential architectures […] artifacts rather than Platonic entities' (1999b: 162). Thus, the HPC view does rule out *Platonic* realism, but may not rule out other forms of realism.

Importantly, the HPC view is, in a sense, neutral as to which properties within the cluster we should most care about when trying to identify instances of a given kind. The literature in the metaphysics of words shows that people have widely differing intuitions about whether semantic, orthographic, phonetic, intentional, or historical properties are the most important when it comes to correctly saying that some token is of a given type.

Adopting the view that words are homeostatic property clusters allows us to accept a form of neutral pluralism about these debates. The reason that intuitions vary is because particular instances of all of these sorts of properties are part of the clusters of properties that are caused by the relevant homeostatic mechanisms. The HPC view can thus defuse this debate. Rather than trying to find which sort of property is the most important we can recognize that all of them arise from stable mechanisms and this explains why they regularly co-occur. None thus have a privileged metaphysical status, though each may have a privileged *epistemological* status depending on what aspects of words we think is most important to our explanatory aims.

More broadly, adopting an HPC view of words also has the benefit of aligning our account of word-kinds with the current dominant views on other biological and psychological kinds (see, for example, Machery 2009, Griffiths 1997, Taylor 2018; 2019). The HPC view has been

widely adopted by philosophers of biology and the social sciences: as Samuels and Ferreira write that 'philosophers of science have, in recent years, reached a consensus—or as close to consensus as philosophers ever get – according to which natural kinds are *Homeostatic Property Clusters*' (2010: 222). The HPC view of words brings our metaphysics of words in line with other kinds posited within similar scientific frameworks.

The balance that the HPC view has been cited to provide between 'natural flexibility and explanatory integrity' (Wilson, Barker & Brigandt 2007) also allows this ontological theory to reflect the science of language more closely. Linguistic theories acknowledge and attempt to reflect the changing nature of their subject matter whilst trying to provide strong explanatorily powerful accounts of how language actually works in the world. The HPC view of words, as in other domains that it has been applied to, provides this balance between these aims also.

Indeed, the HPC view suggests that identifying kinds is a matter for empirical a posteriori inquiry, and not a priori investigation. This is because the identification of clusters of properties is not sufficient for defining a kind; we also need to identify the mechanisms underlying those properties, and identifying these requires empirical rather than a priori methods (Wilson, Barker, and Brigandt 2007). In the case of word-kinds, the methods of various scientific disciplines may be invoked. Linguistics will, naturally, be central. But other sciences that consider the sociological and cognitive bases of language may also play a significant role in the identification of kinds. The HPC view of words encourages closer ties between the metaphysics of words and the empirical study of language.

Lastly, and even more broadly, this discussion may serve as a template for the application of the HPC view to other social entities and kinds. Words are themselves clearly social entities. If the HPC view models words well, then it may also be preferable to essentialism with respect to other types of social entities too.

Acknowledgements

References

Anderson, R. P. and Handley, C. O. 2001. A new species of three-toed sloth (Mammalia: Xenarthra) from Panama, with a review of the genus Bradypus. *Proceedings of the Biological Society of Washington*, 114: 1-33.

Boyd, R. N. 1990. Realism, approximate truth, and philosophical method. In: C.W. Savage (ed.), *Scientific theories*. Minneapolis, MN: University of Minnesota Press. pp. 355–391.

Boyd, R. N. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61: 127–148. DOI: https://doi.org/10.1007/BF00385837.

Boyd, R. N. 1999a. Kinds, complexity and multiple realization. *Philosophical Studies*, 95: 67–98. DOI: https://doi.org/10.1023/A:1004511407133.

Boyd, R. N. 1999b. Homeostasis, species, and higher taxa. In: R. A. Wilson (ed.), *Species: New interdisciplinary essays*. Cambridge, MA: MIT Press. pp. 141–185.

Boyd, R. N. 2000. Kinds as the ''workmanship of men'': Realism, constructivism, and natural kinds. In: J. Nida-Ru¨melin (ed.), *Rationalitat, Realismus, Revision: Vortra¨ge des 3. Internationalen Kongresses der Gesellschaft fur Analytische Philosophie*. Berlin: De Gruyter. pp. 52–89. DOI: https://doi.org/10.1515/9783110805703.52

Boyd, R. N. 2003a. Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, Part I. *Philosophy and Phenomenological Research*, LXVI: 505–553. DOI: https://doi.org/10.1111/j.1933-1592.2003.tb00278.x

Boyd, R. N. 2003b. Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, Part II. *Philosophy and Phenomenological Research*, LXVII: 24–47. DOI: https://doi.org/10.1111/j.1933-1592.2003.tb00024.x

Brandom, R. (1994). *Making It Explicit*. Cambridge, MA: Harvard University Press.

Bromberger, S. 1989. Types and Tokens in Linguistics. In: A. George (ed.), *Reflections on Chomsky.* Basil Blackwell. pp. 58-89.

Burge, T. (1979). 'Individualism and the Mental'. *Midwest Studies In Philosophy*, 4: 73-121. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Cappelen, H. 1999. Intentions in Words. *Nous*, 33(1): 92-102. DOI: https://doi.org/10.1111/0029-4624.00143

Ereshefsky, M. 2010. What's Wrong with the New Biological Essentialism. *Philosophy of Science*, 77: 674–685. DOI: https://doi.org/10.1086/656545

Ereshefsky, M. and Reydon, T.A.C. 2015. Scientific Kinds. *Philosophical Studies*. 172: 969–986. DOI: https://doi.org/10.1007/s11098-014-0301-4

Goodman, N. 1955. *Fact, Fiction, and Forecast*. Cambridge, Massachusetts: Harvard UP

Griffiths, P. E. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press. DOI: https://doi.org/10.7208/chicago/9780226308760.001.0001

Irmak, N. 2018. An Ontology of Words. *Erkenntnis*. 84.5: 1139-1158. DOI: https://doi.org/10.1007/s10670-018-0001-0

Kaplan, D. 1990. Words. *Aristotelian Society Supplementary*. 64: 93–119. DOI: https://doi.org/10.1093/aristoteliansupp/64.1.93

Kaplan, D. 2011. Words on Words. *Journal of Philosophy*. 108 (9): 504-529. DOI: https://doi.org/10.5840/2011108926

Katz, J. J. 2000. *Realistic Rationalism*. Cambridge: MIT Press.

Machery, E. 2009. *Doing Without Concepts*. New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195306880.001.0001

Miller, J.T.M. 2020a. On the Individuation of Words. *Inquiry*. Online first. DOI: https://doi.org/10.1080/0020174X.2018.1562378

Miller, J.T.M. 2020b. The Ontology of Words: Realism, Nominalism, and Eliminativism. *Philosophy Compass*.15(7): 1-13.

Miller, J.T.M. 2021. A Bundle Theory of Words. *Synthese*. 198, 5731–5748. DOI: https://doi.org/10.1007/s11229-019-02430-3

Nefdt, R. M. 2019. The ontology of words: A structural approach. *Inquiry*. 62(8), 877– 911. DOI: https://doi.org/10.1080/0020174X.2018.1562967

Peregrin, J. (2008): 'Inferentialist Approach to Semantics', *Philosophy Compass* 3, 1208-1223. DOI: https://doi.org/10.1111/j.1747-9991.2008.00179.x

Putnam, H. (1975). 'The meaning of 'meaning'', *Minnesota Studies in the Philosophy of Science* 7:131-193.

Rieppel, O. 2010. New Essentialism in Biology. *Philosophy of Science*. 77(5): 662-673. DOI: https://doi.org/10.1086/656539

Samuels, R. and Ferreira, M. 2010. Why don't concepts constitute a natural kind? *Behavioral and Brain Sciences*. 33: 222–223. DOI: https://doi.org/10.1111/phc3.12691

Szabó, Z. 1999. Expressions and their representations. *The Philosophical Quarterly.* 49(195), 145–163. DOI: https://doi.org/10.1111/1467-9213.00134

Taylor, H. 2018. Emotions, concepts and the indeterminacy of natural kinds. *Synthese.* 197(5): 2073-2093. DOI: https://doi.org/10.1007/s11229-018-1783-y

Taylor, H. 2019. Whales, fish and Alaskan bears: interest-relative taxonomy and kind pluralism in biology. *Synthese.* 198(4): 3369-3387. DOI: https://doi.org/10.1007/s11229-019-02284-9

Wetzel, L. 2006. Types and Tokens. *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition). E. N. Zalta (ed.).

Wetzel, L. 2009. *Types and Tokens: An Essay on Abstract Objects*. Boston, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262013017.001.0001

Wilson, R., Barker, M., and Brigandt, I. 2007. When Traditional Essentialism Fails: Biological Natural Kinds. *Philosophical Topics*. 35(1/2): 189-215. DOI: https://doi.org/10.5840/philtopics2007351/29