

On Communication-Based *De Re* Thought, Commitments *De Dicto*, and Word Individuation

Adèle Mercier
Queens University

Preliminaries

In the old days of descriptivism, reference was thought to be a function of ones *epistemic* connection to a referent, a connection fully established by ones beliefs about the referent, which were taken as individuating of the content of ones thought. Thus it was thought that when I think of Aristotle, the object of my thought is whoever instantiates (enough or most or a weighted sum of) the properties that constitute my information about Aristotle. Likewise, it was thought that when I think of water, the object of my thought is whatever satisfies (enough or most or a weighted sum of) the properties of being a colourless, odourless, tasteless, potable liquid originally flowing in rivers and lakes.

Since different people have different epistemic backgrounds and perspectives, it was natural to think, on a view where the connection to referents is epistemic, that different people could attach different senses to their words. And to the extent that the sense of a term is what determines its referent, it was natural to think that different people could mean different referents by the same word, or perhaps stated more precisely, by the same word-*form*. In short, the view was fully supportive of individualism about mental content, and thus about reference.

With Donnellan's *Proper Names and Identifying Descriptions* (1970), Kripke's *Naming and Necessity* (1972) and Putnam's *The Meaning of Meaning* (1975), such views have been --to my mind largely rightly-- debunked in favour of a causal theory of reference. But, like Searle in *Proper Names and Intentionality* (1983), I want to argue that the new theory of reference that has substituted to descriptivism an externalist defense of reference has still not refuted (some aspects of) individualism (though the aspects I have in mind are not wholly Searle's). I wish to draw attention to individualist contributions to the semantics and metasemantics of words that current views fail to acknowledge.

Under the causal theory, the idea that reference is a function of ones epistemic connection to a referent has been replaced by the idea that reference is a function of ones *metaphysical* connection to a referent. What connects me to a referent, on the post-descriptivist view, are actual facts --causal, historical facts-- and not what *I take* those facts to be. To quote Kripke: On our view, it is not how the speaker thinks he got the reference, but the actual chain of communication, which is relevant (p. 93). Since the substance with which I have been

causally interacting when I utter the words this is water happens actually to be H₂O (unbeknownst perhaps to me, and indeed, even if unbeknownst to us all), it is H₂O that is individuating of the content of my thoughts about water. Since the person who was being talked about when I interacted historically with the name Aristotle happened actually to be *Aristotle* (whoever that is), it is *that* person --the person standing at the end of the causal-historical chain that brought the name Aristotle down the generations to me-- who is individuating of the content of my thought about Aristotle. As Kripke puts it: In general, our reference depends not just on what we think ourselves, but on other people in the community, the history of how the name reached one, and things like that. It is by following such a history that one gets to the reference (p.95)

Viewing reference as a metaphysical rather than an epistemic affair guarantees an externalist view of reference, and to the extent that reference is what individualism is about, it constitutes a refutation of individualism. These views are well-worn and commonplace by now, and I shall take them for granted.

Let us now distinguish between the semantics and the metasemantics of words. The semantics of a word is simply what it means, and in the case of a directly referential term, what it refers to. (Indexicals impose qualifications on this simple statement, but let us forget about indexicals here.) According to causal theorists, the semantics of a term is what is passed down the causal chain. Because it is passed down the causal chain, it is not under the control of the user. As Kaplan puts it: We are, for the most part, language *consumers*. Words come to us *pre-packaged* with a semantic value. (Afterthoughts, p. 602)

The metasemantics refers to the event whereby a term is created, before it is launched down the causal chain. A name can be created by at least one of the following two ways: Its reference can be fixed by description (as when Leverrier coins the name Neptune to refer to whatever it is, unbeknownst to him, that is the cause of the perturbations in the orbit of Uranus); or its reference can be fixed by ostension (as when Aristotle's mother points to her baby and baptizes it Aristotle). Other words have their meanings fixed by similar processes, subject to complications which we can ignore for the nonce.

Note that reference-fixing by ostension typically reduces in the end to reference-fixing by description. This is because descriptive information is (however implicitly) needed to individuate demonstrations: Aristotle's mother is baptizing the whole baby in her arms, not its nose, and were she to have inadvertently brought the neighbour's baby to the baptismal ceremony, we could still allow, I think, that she has baptized her son Aristotle (and not re-baptized the neighbour's son). Let's call this the ***Principle of the Description-Ladderness***

of Ostention (the pedantic language is intended to help recall the principle later).

Note also the greater *autonomy* of the individual with respect to the metasemantics than with respect to the semantics of a name. It is obviously not up to Leverrier *what* the cause of the perturbations in the orbit of Uranus is (that's up to God, and for us to find out); so it is not up to Leverrier what *thing* the referent of Neptune is. Nevertheless, it is fully up to Leverrier as the coiner of the term what *sort of word* Neptune is (a proper name for an individual, a kind term, an adjective, and so on). Likewise, though the criteria of transworld identification of Aristotle are not up to Aristotles mother (they depend on such things as the actual egg and sperm out of which he was conceived), it is fully up to her as the coiner of the name Aristotle to determine what sort of name it is (a generic family name, a proper name for a person, for a time-slice of a person, for an undetached person-part, and so on). The case of kind terms (barring complications for now) is similar: if I point to a liquid while coining the word blik, though I am not the expert needed to individuate what the very stuff I've pointed to is (we need a chemist for that), nevertheless I am authoritative about what sort of kind I intend the word to stand for: for example, *this very stuff* (as individuated by our scientific community) and nothing but it, or any liquid relevantly *like* it, or its *color*, or any of its phenomenal properties. My intention will determine what *sort* of word I've coined: a natural kind term, a phenomenal kind term, a color term, and so on. And what sort of word it is will in turn determine, not what thing, but what *sort* of thing its referent can be. More on this later.

Finally, note that one central thesis of the causal-historical picture is that semantic events, or the passing of meaning down the chain, massively outnumber (by orders of magnitude) metasemantic events. And while I think this is indeed largely true, I will try to show that much more metasemantic creationism takes place than is allowed for by the current view (and that is where individualism still has a foothold).

Now, suppose you use the word water and I want to know what you mean by that word. There are two readings of what it is that I might want to know. On one reading --call it the object-level reading-- I want to know what the *stuff* is to which you're referring by that word. The answer to this question depends on what the stuff that you've been interacting with and calling water happens to be. The question can be rephrased in Twin-Earthian language as the question:

Do you mean *this* stuff? (pointing to H₂O) i.e. Do you inhabit the actual world?

Or do you mean *that* stuff? (pointing to XYZ) i.e. Do you inhabit Twin-Earth?

And the lesson to be drawn from the demise of the description theory is that this question is a question for

metaphysics, not epistemology. Its a matter of causal connection *to the world* (and, in the case of causal connections to kinds, its also a matter of causal-historical connection to a science or scientific community individuating those kinds). Most importantly for our purposes, its a question whose answer lies outside the competence of the individual.

Another reading of what it is that I want to know is meta-linguistic. We can put it in terms of the following question:

Which of the many possible homonymous words water are you uttering by means of this token of the word-*form* water?

According to causal theorists, though a different question, this is *also* a question of metaphysics, and *also* a question whose answer lies outside the competence of the individual. Its a question of causal history: In this case, a question about your causal connection *to a language* (or to a community speaking this language). If you speak English, for instance, the word you are uttering is (possibly unbeknownst to you) the word meaning water (whatever substance water may be in the world you happen to inhabit: e.g. H₂O, XYZ, etc.). If you speak French, it is (possibly unbeknownst to you) a different word, namely that word water (pronounced slightly distinctly) which refers to a toilet (whatever toilets may be in the world you happen to inhabit: e.g. bathroom fixtures, religious artefacts, etc.). The choice of homonym available to you depends, as Kaplan might put it, on what word, pre-packaged with what semantic value, is causally available in the language you speak for you to consume.

I shall argue that the second question is a far more nearly epistemological --and individualistic-- question than it is fashionable these days to argue. To motivate the need to revisit the issue, I first examine how people have used the new theory of reference to explain how *de re* thought can be communicated through the medium of language. I think the purported explanation is hardly explanatory, and that its weakness sheds light on problems with the causal-historical picture.

On Communication-Based *De Re* Thought

The standard (easy) case of *de re* thought is when the thinker has encountered the object the thought is about perceptually, or when the thinker has a memory-based thought derived from perception. Several people have raised in recent years the question of whether one can also have *de re* thoughts about things one has *not* encountered, but of which one has merely been informed. Like most such people, I think the answer to this question is that we can (and I will simply be assuming for the sake of the current argument that we can). Unlike

most such people, however, I think the answer that is standardly given about *how* we can, is not satisfactory.

The answer that is standardly given about how we can have *de re* thoughts about particulars such as Aristotle, or about kinds such as African dormice, which we have never encountered (or of which we have never encountered samples in the case of kinds) but of which we have merely been informed, invariably appeals to the causal-historical chain of communication, which is often taken as supplying an account of how language transmits information.

Let me cite a few key passages that I take to be widely representative of the current philosophical orthodoxy on this matter. Noteworthy is Kaplans conception of the role for language in thought, a conception which he amusingly summarizes in his Afterthoughts as *Vocabulary Power as Epistemological Enhancement*.

The notion that a referent can be carried by a name from early past to present suggests that *the language itself* carries meanings, and thus that we can acquire meanings through the instrument of language.

On my view, our connection with a linguistic community in which names and other meaning-bearing elements are passed down to us enables us to entertain thoughts *through the language* that would not otherwise be accessible to us. (p.603)

Contrary to Russell, I think we succeed in thinking about things in the world not only through the mental residue of that which we ourselves experience, but also vicariously, through the symbolic resources that come to us through language.

So how shall I apprehend thee? Let me count the ways. I may apprehend you by (more or less) direct perception. I may apprehend you by memory of (more or less) direct perception. And finally, I may apprehend you through a sign that has been created to signify you. (p. 604)

Now the question of how I can have *de re* thoughts about you merely by apprehending a sign created to signify you can be refocused as the question: What does it take for the linguistic community to transmit *that* sign (i.e. *its* word for you) to me? Kaplan raises this question in his paper called Words. He asks:

Take the utterance or inscription received and the utterance or inscription transmitted. What makes it that the transmission is an utterance or inscription of the *same* word as that received?

And he answers:

The identification of a word uttered or inscribed with one heard or read is not a matter of *resemblance* between the two *physical* embodiments.

[So I can say Adèle Mercier and still transmit my name even if the recipient calls me Aydelee Murciyur (as some unused to French names are wont to do).]

Rather, it is a matter of *intention*: was it *repetition*?

Someone like Burge would make it a matter of intending to *defer*.

Of course, now the question boils down to: Under what conditions is an intention to repeat (or to defer) *successful*?

Note that the answer to this question cannot be, on nothing less than pain of circularity, that the intention to repeat has been successful as long as you've repeated a word *with the same semantic value* as the word you intended to repeat. It's not that this rendition of the success conditions of repetition is inherently wrong (indeed, because I'm not a consumerist, I happen to think that some such rendition is largely right). It's rather that those who are committed to the causal historical chain *as an answer to the second question* --those committed to the view that words come to us pre-packaged with their semantic value-- cannot, on pain of circularity, avail themselves of identity of semantic value as a condition of successful repetition. If the condition of successful repetition is that we be able to look inside you, as it were, (or externally observe your behavior) to see what semantic value your word has, then compare it to the semantic value our word has, and judge them to be the same, then it is not the case that the semantic value of your word is properly parasitic on (or deferential to) the semantic value of our word, in just the way it is purported to be by the historical chain picture of the transmission of semantic values. The historical chain has me committed to the semantic value of the word that is being transmitted to me, whether I know what that value is *or not*. So we can't appeal to what *I take* that value to be (or how I behave with the word) in order to ascertain that a successful repetition has occurred.

But if not by the identity of semantic value between the word emitted and that repeated, how do we ascertain that a successful repetition has taken place? Here I think it will be instructive to review how Kent Bach, in his Thought and Reference, explains the transmission of semantic value from user to user, since no one to my knowledge has attempted to do so with such a degree of exactitude.

A speaker can actually *display* his *de re* way of thinking of the object and thereby enable the hearer to think of it in the same way. If the speaker is thinking of something by name, he is entertaining a mental token of the name; when he refers to it by name, he produces a physical token of that name; and the audience, upon hearing *that* token, forms a mental token of *the same* name. (my emphasis)

Note that Bach is here *assuming* that successful repetition has occurred, not explaining *how* it has occurred, or *how* the mental token formed by the audience ends up being a token of the same *name*, as opposed merely to being a token of the same *name-form*. Though the audience can presume that the speaker has produced a physical token of a name, and not merely a physical token of a name-form, how is the audience to distinguish the physical token of that name, from the physical token of any other name having that name-form, given that all that distinguishes the name from another of the same form is its semantic value (which is not perceptible

and hence not recoverable from the produced token)?

Bach answers this worry thus:

Since the hearers mental token of the name inherits[--this word is in quotes in the original--] the same object as the speakers, the object of the hearers thought is determined relationally.

A token of a name can function as a *de re* mode of presentation because its reference is determined by its ancestry. (p.32)

Note here that so far the relation between the question of *what ancestry the word has* and the question of whether the repetition has been successful, is taken for granted, and left wholly unaccounted for. But here comes the specific proposal:

The token plays this role by being *of a certain form* (sound or shape), generally the same as the one to which it is linked. Indeed, *that only its form matters* is what *constitutes* its being used as a name. And that is what enables one to form *de re* thoughts about an unfamiliar object referred to by that name. Since the token of a name represents *in virtue of its form*, not its meaning, *its representational features can be perceived* by the hearer, who can then and thereafter use mental tokens of the same name to think of (or refer to) the same object. (p.33, my emphasis)

What Bach appears to be saying is that it suffices for repetition to be successful (at least in the case of names) that you have repeated the same *form*. Since the form, unlike the semantic value, can be perceived by the hearer, the hearer is in a position to repeat that form, and to repeat it successfully.

On the Form of Words

Now, of course, as Kaplan has pointed out, we cant make identity of form too strictly a matter of resemblance between the two physical embodiments of the token (we want to allow Aydelee Murciyur to count as having the *same form* as Adèle Mercier). So we cant take the sound or the shape too literally as necessary conditions on the formal identity of the name.

But neither are similarity of sound and shape *sufficient* conditions on formal identity. Support for this claim requires a discussion of what counts as the form of a word. I engage in an all-too-brief discussion of this very vexing question with the use of two cases.

First case:

I claim to know *three* words BARK. I know the *kind* name (or noun) bark, as in what covers the outside of trees. I know the *action* name (or verb) bark, as in the noise emitted by dogs. And I also know a dog

whose *proper* name is Bark. These words may seem formally identical, but I claim that they are not.

In some dialects of English, the words writer and rider are pronounced exactly alike. If speakers of these dialects were illiterate, the words would no doubt seem to them formally identical. Still, they would not be. The word writer is formally related to the verb to write, whereas the word rider is formally related to the verb to ride. The verbs to write and to ride are pronounced differently and entertain distinct relations with other words in the language (the verb to write countenances an indirect object with a *recipient* thematic role: I write a letter to Jones, whereas the verb to ride countenances an indirect object with a *locative* thematic role: I ride a horse to Jonestown); hence they are formally distinct. By transitivity of formal identity (or distinctness), writer and rider are formally distinct.

My three words bark are formally distinct (despite appearances) as witnessed by the difference in their formal relations with other words in the language. It is an accident of writing conventions (admittedly a very convenient one!) that we do not display the whole deep-structure trees of the sentences we write, or that we do not subscript the syntactic category of words as we write them. If we did, their formal distinctness would be obvious. We can think of Latin as a language which does exactly this: it marks nouns with case endings and verbs with declension suffixes, thereby displaying their formal distinctness. This is information that it would be redundant to display in English, precisely because it is recoverable (for the most part) from the formal relations the words entertain with one another in the sentence, such as their word-order (Latin is largely a free word-order language), and their satellites (nouns are accompanied by quantifiers, verbs aren't; verbs require subjects, nouns don't; and so on).

Now I say to you: [bark-is-pretty]. (Pretend this is spoken, or written in the International Phonetic Alphabet, so you don't have the benefit of perceiving whether my statement involves a capital or lower case b.) And you intend to repeat what I said.

You will be able to use your syntactic knowledge to eliminate the action name from the options, because you know that verbs can't be subjects of sentences unless they are infinitives (in which case I would be having you repeat: *To bark* is pretty) or unless they are gerunds (in which case I would be having you repeat: *Barking* is pretty). So you can perceive that I'm not using the action term bark (your syntactic knowledge gives you access to this perception). But there are still 2 options left. Now go ahead: Intend to repeat what I said.

Am I saying that bark is pretty? Or am I saying that Bark is pretty? If I wrote it down, then you could use your knowledge of spelling conventions to perceive which word I'm using, the initial Capital individuating

the word as the proper name, the lack of it as the kind name. Unfortunately, that perceptual access is not available to you here (as it would not be in most cases of linguistic transmissions, which are oral). In your attempt at repetition, are you using the kind name? or the proper name? What *sort* of word are you repeating?

According to Kaplan and Bach, the answer is easy: you are repeating the same sort of word as the word that I said. But this answer is too easy. I claim that you have not repeated my *word*, indeed that you cannot be said to have repeated *any* word, unless you can make a mental commitment as to the sort of word you are repeating (proper name, common noun, verb, and the like). For without thus committing yourself syntactically, you will not be able to use the word productively yourself, to generate sentences using it. You will not know, for instance, whether you ought to say:

I wonder whether *the* bark will still be pretty tomorrow
or
I wonder whether *Bark* will still be pretty tomorrow .

Without such a commitment, what you have repeated is a word-*form*, not a word; you are behaving like a parrot, not engaging in speech. Like a parrot, you will only be able to repeat sentences previously uttered. What you will have learned is a fixed phrase containing the sequence bark (like infants who learn to say Whatsdat but cannot use what and that in other phrases). You will not be a linguistically competent user of any *word* bark.

Now, the point of this exercise is simply to illustrate that not all formal features of a word, indeed not even the most important formal features of a word (such as its syntactic features) are features that the hearer can readily perceive. Note that in the case of repetition of phonetic forms, the speaker has an intuition of formal identity: I say to you: No, not Aydelee Murciyur: A-dèle Mer-cier. And you repeat after me: Ay-de-lee Mur-ci-yur. I say to an Australian: bark. She repeats: baak. Here, somehow, the differing phonological rules of different idiolects or dialects function to establish the membership of these superficially distinct forms into an intuitive equivalence class, in spite of the objective phonetic distance between them. Thats how speakers of distinct dialects manage to understand each other: distinct phones (sounds) are intuited as counting as the same phonemes. But there is more to a words being the word that it is than meets the ear. And some of the important formal features of the word (like the syntactic class to which it belongs) are sometimes (in some contexts) just as unrecoverable as can be its semantic value. Words (at least in English) dont wear their syntactic classes on their sleeves. But in the case of syntactic features, there typically isnt the same kind of intuitive judgment of formal identity as that available for phonological equivalence, to act as a criterion of successful repetition. Different phones can correspond to identical phonemes, but there is nothing (obvious)

akin to that in syntax.

The second case is invoked to show what serious consequences failure to perceive certain features of the word can have on whether the word emitted is successfully being repeated or whether it is not some other event --like an inadvertent new metasemantic event-- that is taking place instead.

Second case:

I witness a commotion involving the police, a citizen and an intellectual-looking book thief. It turns out that the citizen has made a citizen's arrest of the book thief, and the police are discussing the merits of the arrest. The word Aristotle keeps recurring in their conversation. I have never heard that word before but I firmly draw from my experience conclusions that I would voice as that an Aristotle is a form of arrest and that the thief has been Aristotled. I am a properly deferential language user, so the next day I phone my lawyer to inquire what conditions of arrest individuate an Aristotle precisely. My lawyer tells me there is no such thing and that Aristotle was an ancient Greek philosopher. I resolve to consult a more knowledgeable lawyer tomorrow.

Meanwhile, am I having thoughts, even *de dicto*, about Aristotle? Here, as always, intuitions will differ. But I find it *very* counterintuitive to maintain that I am. The word that has (admittedly mistakenly) entered my idiolect has been invested by me with a *syntactic commitment* (as a noun, which morphological rules allow to generate as a verb) that is syntactically incompatible with its being a proper name for an animate object. By the syntactic rules of English as well as of my own idiolect, singular nouns require determiners but proper names for animate objects shun them. If it were suggested to me that I was thinking that Aristotle is a form of arrest, the suggestion would strike me not just as false (given my intuitions), but as grammatically ill-formed. I can no more think *that Aristotle is a form of arrest* than I can think *that door is an opening in the wall*. These thoughts are missing a logical constituent (nothing less than a quantifier). They are incomplete. It is a matter of syntax that I cannot entertain them.

The word that has entered my idiolect has also been invested by me with a *sortal commitment* --as a (would-be) legal kind term-- that is categorially incompatible with its being a proper name for an animate object. I understand the thought that Aristotle is a form of arrest: that is the mistaken thought that the word Aristotle refers in the public language to a form of arrest. But what kind of thought is the thought *de dicto* that Aristotle is a form of arrest? What would it mean to think that there exists something that has both the properties of being Aristotle and a form of arrest? or to think that something is both a form of arrest and a

famous ancient Greek philosopher? What propositions would these sentences express such that one could be related to them in thought? These thoughts are so semantically deviant that I am not sure in what sense one can really be said to entertain them. They are nonsense.

Having checked with a convincing number of lawyers, I finally relinquish my understanding of events. How should I describe my own relinquishment? Two suggestions have philosophical currency.

The causal theorists account goes something like this: I have finally gotten a *better mastery* of the word Aristotle which is the term I have been repeating from the outset; I have finally realized that it is not a common noun but a proper name for an animate object. Whereas before I produced syntactically deviant strings, I now know how to use the word in syntactically well-formed sentences.

The metasemantic creationists account goes something like this: I have discovered that I failed to repeat the word I intended to repeat and instead metasemantically created a new word, the common noun *aristotle* meaning a kind of citizens arrest (with its concomitant verb form yielded by productive morphology). This is a word not shared by others; consequently, I must be very careful when using it to communicate (which means it will soon fade into the dark recesses of unused words). I have also discovered that others use the homonym Aristotle, a proper name for an animate object, and I promptly add this new word to my lexicon, deferring to authorities about its reference.

How shall we adjudicate between these suggestions?

Note that the causal theorists account countenances a single word, of which the speaker develops progressive mastery. The metasemantic creationist account, on the other hand, countenances two words, of which one is more socially useful than the other. Briefly, three reasons are standardly put forward for preferring a one-word account: One, that it is a requirement of commonality of concepts, and hence of communication; two, that without it, we could not explain progressive mastery of a word; three, that word-creation is the privilege only of writers, journalists, and editorialists in Vogue [Kaplan].

I shall not discuss these reasons except to state very briefly why I find none of them convincing. The third is a popularly held normative belief; but who initiates the existence of new words is an entirely empirical matter, better determined by linguists than by the naive populace. (I claim without argument here that the most productive creators of new words are not the literati but bilinguals and children.) The first and second reasons stated are dubious for the following reasons: I have just invented a verb *to aristotle*. It means to perform a kind of citizens arrest. The verb has vague boundaries, because I made it up and I'm vague about citizens arrests.

Many (perhaps most) words have vague boundaries (is wine that has begun turning to vinegar still wine?). But it seems to me that you and I have succeeded in communicating, and in sharing in the concept of an aristotle, as well (or as poorly) as we do the concept of wine. It is true that there has been no progressive mastery of a single word; rather, there has been progressive mastery *of the social language*, in the form of a substitution of a socially useless word in favour of a socially useful one. To insist that what we need is always an account of progressive mastery *of a word* is to beg the question in favour of the one-word approach. It is also to miss the important distinction between acquiring more information *about the referent* of a word, and being so confused *about the word* as not to succeed in acquiring it at all.

There is essentially one reason for preferring a two-word account over a one-word account. But it is a very good reason. It is that the vast majority of our words share a causal history with other, formally distinct, words. For example, the verb to table is formally a different word than the noun table, as witnessed by the fact that they are not substitutable *salva veritate*. In fact, they are not substitutable *salva formulae* (their substitution destroys well-formedness). The words warden and guardian are different words (it is possible to know one without knowing the other) despite the fact that they are just distinct pronunciations of the same original word. The same holds for the English word wasp and the French word guêpe. And in spite of their near perfect resemblance, the English word formidable is a very different word from the French word formidable (as I once discovered after thanking English-speaking hosts for their formidable hospitality). Yet in all these pairs, both words share a causal history --up to their difference, that is. The general problem with the one-word account, or with word-individuation by causal history, is that it yields way too few words (and way too few languages).

On Formal Identity and Word Individuation

The cases that I have provided above argue to the conclusion that the phonetic or orthographic form of words does not suffice to establish their formal identity. Two words can sound (or be spelled) exactly alike and yet be formally distinct. I now turn to argue that even formal identity is not sufficient to individuate words. Two words can be formally identical and yet be distinct words. The argument begins with a tale of two countries.

Marco Polo is shipwrecked on land that he has explored and knows to be an island. (Liberties with historical detail have been taken in service of the argument.) He utterly ignores the existence of the African

mainland, and, let us say, given the geographic information gathered during his travels, he would vehemently deny its existence. So Polo has the belief: There is no mainland within proximity of this island. A native from a country on the African mainland --a country called Madagascar (allegedly from Mogadiscio)-- appears (one who felicitously speaks English) and the following happens: Polo introduces himself as coming from Venice and as having survived the shipwreck of the Santa Venezia; mistakenly assuming that the native lives on the island, he asks her what the name of her country is, to which she replies Madagascar.

At least a new *word-form* has been transmitted from the natives language to Polos. Polo has the general semantic intention typical of language learners: he forms a deferential metalinguistic intention to use the word Madagascar to stand for whatever the native uses it to stand for. At the same time, Polo assigns certain sortal features to the new word in his lexicon. This is something he must do to be able to use the new word-form productively in syntactically well-formed strings. As it happens, the semantic and syntactic contexts in which the question is raised and answered felicitously allow Polo to perceive correctly at least some sortal features of the word-form --namely, that it is a *proper* name, and a proper name for a *land mass*.

But they might not have. The context might have (mis)led Polo to perceive the name Madagascar as a proper name for a ship, say. For instance, he could have (mis)heard the natives hesitant euh before Madagascar as the Madagascar and inferred that the native had (mis)heard his word country as dory. Polo would then have produced syntactically deviant sentences (inserting the determiner the before the name), or more precisely, sentences that would have seemed syntactically deviant to the native while seeming well-formed to Polo. This ought to suffice as proof that Polos deferential intention is so easily defeasible, that his linguistic behavior cannot be individuated by it: if the word we should countenance Polo as using is determined by his deferential intention, then the syntactic rules of Polos idiolect have to be described as allowing a determiner before a (singular) land mass name, something which (by hypothesis, as in English) they do not allow. But let me not press the point.

Happily, as it happens, Polo has assigned to the word the same formal features as has the native. So he is able to communicate with the native with sentences that seem to both to be syntactically well-formed and semantically non-deviant, like: When I return home, I intend to tell my people about Madagascar.

The fact that Polo is capable of producing well-formed sentences using the new word shows that he has acquired a new word, though it does not settle the question of *which* new word it is that he has acquired. And here the causal theory is led into a dilemma. If Polo has acquired a different word, a new word Madagascar

applying to the island on which he is stranded --call it Madagascar₂-- then the causal theorist is left with the question of how he could have consumed *that* word, given that it is nowhere around to be consumed. If he has repeated the same word as the native, then his intention *de dicto* is to tell his people about the natives mainland country. But this intention, even *de dicto*, is quite dubiously attributable to Polo. How can he intend to inform his people of something he not only knows nothing about it but whose existence he would vehemently deny? Moreover, the view that he has received *the* word that was being passed down along the chain --call it Madagascar₁ -- is inconsistent with our feeling that the word he has passed down to *us* is Madagascar₂, the word whose semantic value is the island, not the one whose semantic value is the country on the mainland.

Polo begins his travel memoirs with the sentence: Madagascar is an island. This sentence expresses a false belief in Polos idiolect, if the causal-historical picture is right. It expresses a true belief only if he has single-handedly invented the new word Madagascar₂ and is using that word in that sentence. Had Polo read his memoirs to the native, the original linguistic mistake might have been discovered, and hence corrected, and Polo would have changed the mode of expression of his belief. But he didnt.

Two unfortunate consequences follow from the view that the word picked up by Polo is to be individuated by its causal antecedents, i.e. that he has actually picked up the word Madagascar₁. The first

results from Bachs view of communication-based *de re* thought. Paraphrasing Bach to fit our case:

The native has actually displayed her *de re* way of thinking of the object and thereby enabled Polo to think of it in the same way. The native was thinking of something (the country on the mainland) by name, and thus entertaining a mental token of the name; when the native referred to it by name, she produced a physical token of that name; and Polo, upon hearing that token, formed a mental token of the same name. Since the token of the name represents in virtue of its form, not its meaning, and since its representational features were perceived by Polo, he can then and thereafter use mental tokens of the same name to think of (or refer to) the same object (the country on the mainland).

This is indeed a puzzling view. We began our story with Polo vehemently convinced that he was surrounded by nothing but water. He picks up a word from some perfect stranger, and with no new information whatsoever, Ta-Da!, he is not merely (unbeknownst to himself) *referring* to some country on the mainland, but actually able to have a *de re thought* about it. (Take comfort in this story next time youre stranded on some lost island.)

The second unfortunate consequence of the causal-historical picture results from the predictions of the view about what happens when Polo comes home and tells us that he has been stranded on the island of

Madagascar. By the historical chain of transmission, Polo is talking about the mainland. So when he says the island of Madagascar he is using a semantically anomalous construction, on a par, say, with the island of Idaho. But since we cannot (and, according to the theory, need not) perceive the semantic value of the term, we are oblivious to the semantic anomaly and we promptly form a belief that we express as: Madagascar is an island. This is a false belief given our intention to repeat the same name as Polo was intending to repeat, and given that the semantic value of *that* word is a country on the mainland (Madagascar Mogadiscio), not an island. So, courtesy of the causal-historical chain, we are all mistaken (even the Malagasy are mistaken!), that Madagascar is an island.

Worse than that, since in virtually all cases we ignore the causal origins of the actual chains of communication linking us to the words we use, in all good conscience, we must be sceptical of virtually *all* sortal claims we typically take very much for granted: That Aristotle was a person, that Canada is a country, that Idaho is not a sort of animal, etc. Since the semantic value of these words is imperceptible but comes pre-packaged with the word down the chain, we have *no* way of checking that the word has not strayed off course. We are at the mercy of lexicography.

This is a reductio of the causal-historical picture, one which Searle has forcefully rendered in his inimitable style, by rhetorically asking if we are ready to accept that our name Plato might turn out to refer to a barstool in Hoboken, New Jersey. Some causal theorists, for instance Kripke, have shown the good sense to use this picture with moderation, witness the presence in Kripkes writings of cautious --though undeveloped-- passages such as this:

The name is passed on from link to link. But of course *not every sort* of causal chain reaching from me to a certain man will do for me to make a reference. There may be a causal chain from our use of the term Santa Claus to a certain historical saint, but still the children, when they use this, by this time *probably* do not refer to that saint. (p.93, my emphasis)

And about Madagascar, Kripke has this to say:

Evans has pointed out that similar cases of reference shifts arise where the shift is not from a real entity to a fictional one, but from one real entity to another of the same kind. [...] Today the usage of the name as a name for an island has become so widespread that it surely overrides any historical connection with the native name. (p.163)

In an effort too counter the reductio, we now pursue the nature and justification of the probably in the Santa Claus quote, and the means available to the causal-historical picture as explanations of *how* the historical connections *can* be overridden. Continues Kripke:

So other conditions must be satisfied in order to make this into a rigorous theory of reference. I dont know that Im going to do this because [...] Im sort of too lazy. (p.93)

I want to point out that some of the conditions that appear to be necessary to make the picture work (all of which involve some measure of individual metasemantic autonomy) are in direct conflict with some of the core tenets of the very picture (consumerism) of which they appear to be necessary conditions.

On Conditions of Existence and Word Individuation

Kripke argues in Naming & Necessity for a principle that we can call the Principle that Existence Precedes Essence. It is a common intuition that, though unicorns (and Santa Claus) don't exist, they could exist; the intuition is that it is simply a contingent fact that white horses with horns don't roam our prairies, and that no person satisfies a weighted sum of the properties from among the cluster of properties attributed to Santa Claus; there is no reason to think that such creatures, or that person, don't exist in other possible worlds.

Kripke argues against this common intuition:

Even if archeologists or geologists were to discover tomorrow some fossils conclusively showing the existence of animals in the past satisfying everything we know about unicorns [...], that would not show that there were unicorns. (p.24)

His point is easy, and, I think, persuasive. In order for me to have modal properties (the possibility of being other than I am in another possible world, as it were), it is necessary that I maintain my identity across possibilities (that *I* be in those other worlds). What makes me *me*, as opposed to anyone else whose merely superficially indistinguishable from me, are my conditions of existence (for example, that I am the product of a certain specific egg and a certain specific sperm). To locate me in another world, you have to find someone in that other world who satisfies my conditions of existence.

Conditions of existence are a function of actuality. We start with the object as given and derive from it the conditions failing which *that* object would not be in existence. In the case of unicorns, if we start with unicorns as they are given to us, namely *as fictitious objects*, we see that it is essential to their individuation that they *have* no conditions of existence. (Or, if you prefer a Meinongian approach, it is part of their conditions of existence as they are given to us that they be fictitious objects.) So the white horned horses in other worlds, though admittedly very unicorn-like, are not our unicorns. What about the creatures discovered by the archeologists? Well, for exactly the same reasons, neither are they. They may be completely unicorn-like, but our unicorns are fictitious objects, and as such they don't leave bones lying around; so whatever bones are being found, they are not unicorn bones.

Note here that Kripke might have expressed his views about unicorns more cautiously. Since he does

not have access to the *actual* chain of communication linking his use of the term unicorn to past uses of that term, *for all he knows*, the word may well have come down to him through a historical chain linking his current use of the term to an original metasemantic event that involved ostension to a certain kind, the bones of which are only now being recovered by archeologists. We might all (except presumably the original coiner of the term) have been wrong all along that unicorns never existed --in which case, of course, the bones uncovered would be unicorn bones. So Kripke might have said, more cautiously, that provided that the actual chain of communication through which the word unicorn was transmitted to him did not tie his current usage to past existing creatures, whatever the archeologists discover are not unicorns.

That is not what he says, though it is probably what he means. He is no doubt assuming that we are right in thinking that the chain of communication leads back only as far as properly fictitious objects. On the other hand, given what he says about Santa Claus --where he acknowledges that children probably mean a fictitious object *even if* we know there to be a chain leading back to a previously existing person-- we can interpret him as saying about unicorns, that even if there is a chain leading back to some previously existing kind, *somehow* we have (probably?) gotten freed from that chain. Otherwise, given that causal-historical word transmission chains leave precious few fossil traces (and none but overly speculative ones beyond a certain past), why wouldnt the discovery of unicorn-like bones strongly suggest to us that the chain *did* indeed lead back to existing unicorns?

In a similar vein, with respect to our use of the word Madagascar, Kripke might say that though there *is* a chain leading back to the country on the mainland, somehow we have gotten disconnected from that chain. Kripke says:

In all cases [where a reference shift occurs], a present intention to refer to a given entity (or to refer fictionally) overrides the original intention to preserve reference in the historical chain of transmission. (p.163)

He continues: The matter deserves extended discussion, a discussion to which he never returns.

There appears to be a real tension between the commitments of the causal-historical picture --namely, that reference is a function of the actual chain of communication, a metaphysical, not an epistemic, fact-- and the existence of reference shifts --which are imbued with epistemology. The kind of (aborted) reference shift I initiated with Aristotle, like the kind of (sustained) reference shift initiated by the namers of Bark, depended on what I shall call *commitments de dicto*. I call them thus to emphasize firstly, that minimal competence with a word requires a mental attitude *about the word* in the form of an assignment to the word of syntactic and sortal

categories, and secondly (and consequently) that there are virtually no thoughts that are *purely de dicto*, in the sense of being purely parasitic on the language with no psychological contribution whatsoever by the language user. It is my understanding of the word as being of a certain sort (or the will of the namers of Bark to make a verb into a proper name) that resulted in a shift in the sorts of things that could be its reference. The reference shift in Polos case depended not on his different understanding of the sort of word it was but on his commitment to use it to refer to the island on which he was standing (on his present intention to refer to a given entity). Epistemic considerations pervade both cases. My reference shift is due to my beliefs about the words syntax; it follows from my interpretation of the syntactic intentions of its users. Their actual syntactic intentions are irrelevant to the word I would have passed down to my children had I not noticed and repented about the errors of my ways. Polos reference shift is due to his beliefs about the words semantics; it follows from his interpretation of the natives referential intentions. Their actual referential intentions are irrelevant to the word Polo passed down to us. If language users intentions to preserve reference (or to defer reference) were truly as resilient as they would have to be for the historical chain picture to provide an explanatory account of word transmission, it is a mystery how an individuals spontaneous commitments *de dicto* (sortal assignment or intention to refer) could override them.

On Intentions to Refer and Word Transmission

Kripkes only attempt to resolve this tension (that I have been able to locate) occurs in the following passage:

The phenomenon [of reference shift] is perhaps roughly explicable in terms of the predominantly social character of the use of proper names [...]: we use names to communicate with other speakers in a common language. This character dictates ordinarily that a speaker intend to use a name the same way as it was transmitted to him; but in the Madagascar case this *social* character dictates that the present intention to refer to an island overrides the *distant* link to native usage. (p.163, my emphasis)

This passage is far from explanatory. The two facts that need reconciling are, first, the fact that *Polo* intends to use the name Madagascar the same way as it was transmitted to him (and hence by the causal-historical picture refers to the country on the mainland), and second, the fact that he manages, somehow, to bring home a word that refers to the island instead (in violation of the causal-historical picture). The appeal to the social character of language is singularly unhelpful. It misses the point that our current --social, if you like-- intention to refer to an island by our use of the word Madagascar **would not exist** if Polo had not brought home a word Madagascar referring to an island. For had Polo brought home the original word Madagascar, we too, by virtue

of our commitment to the historical chain, would be referring by its means to the mainland. For we are not, and only Polo is, epistemically close enough to the island to form an overriding present intention to refer to a given entity.

What really needs explaining is the fact that *Polo* succeeded in overriding the link to native usage, a link that was *not* distant from him but singularly close. Since by hypothesis Polo is alone on the island, the appeal to the social character of language is spurious here.

Even the much bandied about distinction between speaker-meaning and semantic-meaning will not help to elucidate matters here. Say if you will that Madagascar-the-island is Polos speaker-meaning, and that Madagascar-the-mainland is the semantic-meaning. What needs explaining then is how Polos speaker-meaning becomes our semantic-meaning.

The problem is that what gets passed down causal-historical chains are semantic-meanings, not speaker-meanings. We dont want to end up saying that, though Madagascar-the-mainland is really our current semantic-referent, we are always, all of us, using Polos speaker-referent when we speak of Madagascar and thats how we succeed in referring to the island. But if there has been no shift in semantic-meaning from the natives language to Polos language, then the semantic-referent we have inherited from Polo that has been passed down to us by the chain remains Madagascar-the-mainland.

The problem boils down to explaining how Polos speaker-referent becomes his semantic-referent, so that it can become our semantic-referent. (It doesnt actually matter where along the chain, if not with Polo, the change occurs; the fact is that the speaker-referent of some consumer or other along the chain has to become a semantic-referent so that it can become our semantic-referent.) But the appeal to the social character of language will not explain how Polos speaker-referent (or anyone elses) becomes his semantic-referent. The social character of language could possibly play a part in such an explanation only if there existed at some point a social language in which the semantic-referent of Madagascar were Madagascar-the-island. But the problem is that the causal-historical picture itself precludes the generation of such a language, by precluding the generation of individual speakers who could together constitute the speakers of such a language. Simply to stipulate that a new semantic referent occurs when a sufficient number of speakers share new speaker-referents is not to explain it.

To begin to explain how Polos speaker-referent can become his semantic-referent, I claim, we have to look at *individualistic* aspects of Polos interaction with the *word* Madagascar as it gets transmitted to him.

Alternatively or additionally, to begin to explain how the speaker-referents of each of the members of the critical mass necessary to bring about a semantic change in the social language can become the semantic referents of each of them, for all the same reasons, we have to look at *individualistic* aspects of each of *their* interactions with the word Madagascar as it gets transmitted to each of them as links in the chain. These individualistic aspects can *change* the word (that is, can subvert the chain), thereby changing its semantic referent. Conversely, where the chain is not broken but succeeds in handing the same word down the generations, it is *because* each individual, as a link in the chain, has individually succeeded in reproducing the same word. Historical-causal chains cannot provide an explanation of word transmission because they exist only in retrospect.

On Commitments *de dicto* and Conditions of Existence

It is in our individualistic interaction with words that commitments *de dicto* play a crucial role. To illustrate how they do, let me start with the coining of a new word. Whenever a new word is coined, whether its reference is fixed by ostension or by description, the coiner is always, however implicitly, investing the word with sortal commitments, that is, commitments as to the sort of word the new word is. When the reference is fixed by ostension, there is always an implicit sort of thing that is being demonstrated, which correlates with the sortal commitment of the word that is being coined (recall the pedantically called Principle of Description-Ladenness of Ostention). The same goes where the reference is fixed by description. Though Leverrier does not know what thing is causing the perturbations in the orbit of Uranus (he doesn't know if it's a planet, a star, a meteorite shower, a god), as I claimed above, there have to be some sortal conditions on his succeeding to refer. Let me now illustrate why this is so.

Suppose that the universe is very unstable, and every second a new object comes to perturb the orbit of Uranus. By coining the name Neptune for the cause of the perturbations in the orbit of Uranus, has Leverrier thereby named Neptune the *series* of such causes? Is *that* the sort of thing that can be satisfied by his description the cause of the perturbations in the orbit of Uranus? Or is it rather that, on this supposition, Leverrier has simply failed to refer to anything on the grounds that a series of causes is not a particular and hence the description fails to be uniquely satisfied? Suppose Leverrier has miscalculated the orbit of Uranus, and there are no perturbations. Is Neptune just another name Leverrier has invented for Nothing, or is it rather a name he has invented for nothing?

We may not always be in a position to know the answer to such questions. But I maintain that Leverrier *would* --indeed, would *have to*-- have some intuitions (if probed) about the sortal conditions that must be met in order for him to succeed in referring to something (and, complementarily, intuitions about the conditions failing which his efforts constitute an abortive attempt at naming). For without such sortal intuitions, Leverrier would not know whether he has coined a proper name, a kind term, or some term of another category, information he needs if he is to use the word productively, that is, in novel sentences. For Leverrier to have such intuitions is tantamount to his having intuitions about what he would have to learn about how the world is to conclude that he had not succeeded in referring at all by means of the name Neptune.

Even when we do not coin a word ourselves but rather receive it allegedly pre-coined (or semantically pre-packaged as Kaplan would have it), again, a condition of our being able to use the new word productively is that we commit it to being a word of a certain sort. It will often, perhaps typically, be the case that the preexisting sortal commitments of a word being transmitted down the chain will be recoverable from the sentential context in which the word first appears to us, as when Polo tells us that he has just returned from Madagascar. Since we understand that one returns from a place, and that since Polo is an explorer, the sort of place he has in mind is probably a land mass, our sortal assignment will likely reproduce that of the transmitted word. Madagascar will enter our lexicon as a land mass term. We will thus have repeated a word that is formally identical to Polos, that is, one whose occurrence in a sentence satisfies the same syntactic and categorial restrictions.

Moreover, since we are deferential about the semantic value of this proper land mass name --*and* provided that we do not form a present intention to refer to a given entity-- we will thus have repeated a word that is semantically identical to Polos. We will *consequently* have added a new link in the ongoing chain.

But sortals are often not perceptible. Different sortal kinds can occur in the same syntactic contexts. For example, a sentences such as:

I cant wait to return to Madagascar

is consistent with Madagascar being a persons name as much as a place name. And a language (such as Catalan for example) which put determiners in front of proper personal names makes using the syntax to distinguish proper names from common nouns unavailable, as in:

Im relieved to have found the John.

In such cases, the hearer may assign to the transmitted word different sortal commitments than those it had in

the mouth of the speaker, thereby unwittingly performing a metasemantic act initiating a formally distinct word, that is, one whose occurrence in sentences other than those from which it was transmitted will be subject to distinct syntactic or categorial restrictions.

The exact conditions under which the context of acquisition of a word will maintain or change its sortal commitments may be too complex and aleatory to specify. But I suggest the following as a test for ones *de dicto* commitments. We can call it the Conditions of Existence Test. It consists in consulting ones intuitions about how the world would have to be in order for one to conclude that a word as one uses it is an abortive attempt at referring (even in fiction). The properties one intuitively will not suffice to determine the referents of these words. Nor should they, since we do not always know enough about the referents of our words to identify them. But they do constrain the sorts of objects the referents must be. It is indeed because our (current) word unicorn is the sort of word it is that it cannot have as its referent the sort of thing the archeologists are uncovering. It is, in turn, that sort of word (a word assigned to fictitious discourse) not because of metaphysical facts about the actual chain of communication, since for all we know that chain does reach back to existing objects. Rather, it is that sort of word because of some aspect of our epistemic state as the word became fossilized into our lexicon. The relevant aspect of this epistemic state consists in whatever explains our *de facto* assignments of commitments *de dicto*. An explanation of these assignments can only be made on a case by case basis, if at all; but all my argument requires is that we do make them, and I think I have shown why we must. Nothing less than our ability to use words productively depends on it. And if productive use is not the core of word competence, I don't know what could be.

The case of Polo is different from simply deferential cases because of his present intention to refer to the given island on which he is standing. But it is wrong to think simply that Polo's intention to refer to the island is an overriding intention. Polo, after all, does intend to defer, and he would have relinquished his use of the word had he been corrected in a timely fashion. At least at first, he intends to defer more than he intends to refer: he is no Humpty Dumpty. But in the absence of a correction, it is his *belief* that the deferential intention is satisfied which allows the referential intention to take over. It is this belief which is playing the causal role in his assignment of semantic value to the word Madagascar, that is, in individuating as new the word Madagascar making its way into his idiolect.

Semantic change is as pervasive a phenomenon as it is because the conflict between referential and deferential intentions is often too subtle for detection. Because linguistic contexts are underspecified, people can

easily be (mis)led to believe that their deferential intention has been satisfied, and so the fact that the referential intention has overridden the deferential intention escapes correction. That is how a hearer may reproduce a word formally identical to that transmitted, yet still create a substantively new word.

On Conditions of Existence and the Contingent A Priori

It is a contingent fact about Neptune that it ever caused perturbations in the orbit of Uranus. But this (if Kripke is right) is a fact that Leverrier knows *a priori*. He knows this *a priori* precisely because he has used the description to fix the referent of Neptune to whatever satisfies it. He knows it *a priori* in the sense that it is knowledge (in whatever sense that we can call it knowledge) that he has gotten simply by his understanding of his own language. Likewise, it is a contingent fact that Canada is a land mass on earth; but in exactly the same way that it makes sense to say of Leverrier that he knows *a priori* that Neptune is the cause of the perturbations, I think this is a fact that I know (in whatever sense that we can call it knowledge) *a priori*: not because that I am innately endowed with geographic information, but because the word Canada has entered my lexicon *as* a land mass name. Perhaps I am wrong about this --in which case I will have to sophisticate my intuitions about myself-- but I cannot imagine any way the world might be where the referent of my word Canada were not a land mass (though it could conceivably be a different land mass than the one I would identify). There may well have been a time when *some* word Canada did not refer to a land mass (but to a people, or a river, or what have you); but *that* word Canada, though causally connected to my word Canada, was never part of *my* language. If I am being duped by an evil genie into believing that there is actually a land mass under my feet when in fact there are no land masses anywhere, then regardless of what Canada means in the mouths of other people, I have failed by its means to refer.

Conclusion

As the cases of bark and aristotle show, the phonological (or orthographic) form of a word is not alone individuating of its formal identity. As the case of Madagascar shows, even formal identity is not sufficient to individuate words. We deny this at the cost of denying ourselves an explanation of language variety and semantic change.

Words do not reproduce as if by mitosis, multiplying exact copies of each other. Their reproduction rather mimicks genetic reproduction: words are complex, multi-layered and deep-structured objects, some of whose features get passed on from parent-word to child-word, some not. This should not be a surprise. Since

words have to be absorbed by minds in order to be productively used by them, the copying-fidelity of word-transmission can only be as accurate (or as inaccurate) as the copying-fidelity of mind-reproduction. Our words are genetically different because our minds are.

The causal-historical picture has the cart leading the horse. Polos attempt at repetition of the very word Madagascar which was being transmitted to him by the native has failed. It has failed because Polo inadvertently assigned a different semantic value to the word. *As a consequence*, the historical chain of communication was broken, and a new word came into being. By contrast, our attempt at repetition of the word Madagascar that was transmitted to us by Polo has succeeded. It has succeeded *because* we managed to reproduce Polos commitments *de dicto*. It is a consequence of this fact that the historical chain of communication remained unbroken, and that the same word was transmitted.

Criticisms of the description theory have shown (to my mind, roughly conclusively) that, at least where we form no intention to refer to a given entity, we do not assign semantic values to our words. But we can use words productively in well-formed sentences even when we do not know their semantic values. However, we cannot use words productively in well-formed (or semantically non-deviant) sentences unless we think of those words as words of a certain syntactic and categorial sort. And what sort of word a word is will affect --not what *thing*, but-- what *sort* of thing it can have as a referent.

Criticisms of the description theory have not shown that sortal assignment is not a matter of individualistic competence. If we inherited through the causal-historical chain the sortal commitment of the word along with its semantic value, we could not account for how real reference can shift to another real reference, fictional reference can shift to real, and real to fictional. It is plausible to suppose that it is when people started attributing outlandish supernatural properties to Saint Nicholas that they distorted the semantic context of transmission of the name, which encouraged the recipients to adjust their sortal commitments, thereby transforming the name Santa Claus into the name of a fictitious entity.

Metasemantic events are not the sole privilege of original coiners of words. They reflect the syntactic and semantic adjustments that we each make as we acquire words in our lexicon, without which languages would not be in the constant state of drift in which they inexorably are. Far from showing that reference is strictly a metaphysical matter, where (as Kripke says) only the actual chain of communication is relevant, these adjustments in our commitments *de dicto* are imbued with epistemological content. Though our beliefs do not *determine* the referents of our words (it was the mistake of descriptivism to purport that they did), they do

affect the sorts of words we countenance. And that affects, not their referents directly, but the sorts of things that can be their referents. Commitments *de dicto* reintroduce a measure of individual autonomy, which disrupts the flow (and much of the relevance) of the historical chain.

I apologize for the tortuous route of this argument which I now bring to a close, but I hope that through it you see that Kripkes metaphysical principle that Existence Precedes Essence has a metasemantic (read: epistemic) analog. It is that Commitments *de dicto* Precede Reference.