WORDS

David Kaplan

Preface. In the Summer of 1984 I was playing with some techniques for indicating scope, hoping to find a new way into the vexing problem of substitution on directly referential expressions in belief contexts. After chasing Woozles for a week or so I realized that insofar as the problem of substitution in belief contexts was concerned, I had ended up right where I had begun, but under a clutter of technology. At that time I wrote in large letters on a lined yellow pad, "A NEW IDEA IS NEEDED".

It was syntactical technology that I had been thinking about, and it occurred to me that perhaps syntax, in some vague sense, was a key to the puzzles that I had been unable to solve. I ventured that thought, as it applies to single words, in 'Afterthoughts':

> Lately, I have been thinking that it may be a mistake to follow Frege in trying to account for differences in cognitive values strictly in terms of *semantic* values. Can distinctions in cognitive value be made in terms of the message without taking account of the medium? Or does the medium play a central role? On my view, the message—the *content*—of a proper name is just the referent. But the *medium* is the name itself. There are *linguistic* differences between "Hesperus" and "Phosphorus" even if there are no *semantic* differences. Note also that the syntactic properties of "Hesperus" and "Phosphorus", for example their distinctness as *words*, are surer components of cognition than any purported semantic values, whether objectual or descriptional.[1]

Could it be that the elusive cognitive difference between believing that Hesperus is Hesperus and believing that Hesperus is Phosperus rests on nothing more than syntax?

In speaking of *syntactical*, or *syntactical-lexical*, or what, in

[1] Pages 598–599 in 'Afterthoughts' in Almog, Perry, and Wettstein (editors), *Themes From Kaplan* (Oxford University Press 1989).

Demonstratives, I called *presemantical*[2] differences, I am searching for a term to capture the elements of form that are independent of semantics.[3] The awkward, *syntactical-lexical form*, may best convey my meaning.[4]

The question of whether two sentences have the same syntactical-lexical form may seem utterly trivial, but it isn't. It is part of the purpose of this essay to interest you in that question. For example, Saul Kripke's intriguing puzzle[5] about Peter—the man who first heard Paderewski's name used in connection with his musical accomplishments and later heard the name used in connection with Paderewski's political accomplishments and then concluded that these were two different 'Paderewski's—seems to me to involve this sort of *syntactical-lexical form* issue.

My speculations led me to conclude that I had to go back to basics and rethink not just the semantics of names, but their very syntax, the metaphysics of words: How should words be individuated? What is the nature of a word?

These musings eventuated in my Gilbert Ryle lectures, *Word and Belief*, delivered at beautiful Trent University in Peterborough, Canada, and at various other institutions. The content of the present paper is drawn from the middle lecture. The whole *Word and Belief* project is quite speculative. This is the part in which I have the most confidence.

In the first lecture ("A Puzzle About Direct Reference"), I take up the problem posed to direct reference theory by substitution on names within the context of a propositional attitude report. The problem is that according to direct reference theory two names that name the same thing have the same semantic value, and therefore, substituting one for another, even within a belief-context, should not affect the semantic value of the sentence. After arguing for years,

[2] Page 559 in *Demonstratives* op. cit.

[3] Linguists seem to think that "John admires John" and "John admires Jane" have the same *syntactical form* and differ only in what they call "lexicalization" (though they claim that "John admires himself" differs from both syntactically). Logicians seem to think that "John walks" and "Jane runs" have the same *logical form* (except perhaps when each is conjoined with "John does not walk").

[4] When I use the word "syntax" or "syntactical" (sometimes in scare quotes) or "logical syntax" it is usually this *syntactical-lexical form* notion that I have in mind.

[5] In his 'A Puzzle About Belief' in A. Margalit (editor), *Meaning and Use* (D. Reidel 1979)

unconvincingly, that semantic value (properly understood) is not affected by substitution, I hit upon a brilliant, new, and completely successful, strategy: argue, instead, that semantic value *is* affected by substitution. But I also argue that, contrary to my own long-standing misapprehension, this result is not contrary to direct reference theory.

The detailed argument involves an attack on a version of Frege's principle of compositionality, focussing again on the *syntactical* differences, i.e. differences as *syntactical objects*, between sentences of the form "$a = a$" and sentences of the form "$a = b$".[6] This raises the issue: What determines an utterance to have the form "$a = a$" as opposed to the form "$a = b$"? And this issue takes us directly to matters of word individuation.

I am convinced that we can achieve a highly salutary clearing of the air about the nature of language, and especially about some critical differences between natural languages and logicians' idealizations, if we study the ontology and individuation of words.

The present paper is organised into three main parts. In the first part, *What is a Word?*, I contrast two theories of the individuation (one might say, *metaphysics*) of words: the conventional token/type theory, which I call the *orthographic* conception, and an alternative theory, which I call the *common currency* conception. I will try slowly to entice you into abandoning the traditional favorite in favor of the new conception.[7]

The second part focuses on names, distinguishes such words from their genera, and combats certain wrong construals of my views. I also claim to show in this part that there can be cases of distinct words which are both homophones and homographs—I call them phonographs, they sound the same and they are

[6] I have come to think that two sentences whose syntax—perhaps here I should say, whose *logical syntax*—differs as much as "$a = a$" differs from "$a = b$" should never be regarded as having the same semantic value (expressing the same proposition), regardless of the semantic values of the individual lexical items "a" and "b".

[7] In the philosophical tradition, the main idea of the common currency conception may be thought to be implicit in the work of Kripke, Donnellan, and perhaps myself and others, but in work that was done for semantical purposes, not for syntactical purposes. In my view, these particular ideas of recent theories of reference have simply not been sufficiently exploited. I hope that those who wish to encourage a view of the objectivity of language will welcome a useful ally in my analysis of words.

spelled the same—distinct words that are phonographs and whose semantic values are also exactly the same. The semantics is the same, the spelling is the same, the pronunciation is the same, but they are two different words! I will try to show that this odd result is perfectly natural on a *correct* theory of word individuation.

I

WHAT IS A WORD?

1.1 Expressions and their Occurrences. The question of how to individuate linguistic expressions is a delicate one. Here is a famous sentence from the section 'Use Versus Mention' in Quine's *Mathematical Logic*—one of our sacred texts. Quine is talking about expressions and how we refer to them.

> To mention Boston we use "Boston" or a synonym, and to mention "Boston" we use "'Boston'" or a synonym. "'Boston'" contains six letters and just one pair of quotation marks; "Boston" contains six letters and no quotation marks; and Boston contains some 900,000 people.

I would have thought that a logician like Quine, who is used to distinguishing variables from their occurrences, would have immediately seen that "Boston" contains six letters is false. I only count five letters in the name "Boston": a "B", an "o", an "s", a "t" and an "n". There are, of course, two occurrences of the letter "o".

When I utter "Help!", no one thinks that my utterance *is* the word "help". That would give the language too many words. When I utter "Help! Help!", I haven't uttered two words, I've uttered one word twice. Two utterances, one word. Similarly, when I write "Quiet", my inscription isn't identical with the word, rather it is *an* inscription *of* the word. A single word can, and typically will, have many utterances and inscriptions. Uttering and inscribing, or writing, are actions whereby we produce certain concrete, as opposed to abstract, physical objects: utterances (sounds), and inscriptions. These objects, the

utterances and inscriptions, are the physical media by which we transmit words from one to another.[8]

1.2 Interpersonal Transmission. Every word that we know is either one we invented or one that was transmitted to us in one of these ways, by way of an inscription or an utterance. (I haven't resolved how to think about the interpersonal transmission of words through electronic media and through the use of Morse code and so on. Let's not worry about that.) The central claim here is this, that in interpersonal linguistic communication, *inter*personal—I'm very focused here on the distinction between *intra*personal links, and *inter*personal links—in *inter*personal linguistic communication, words must take on a physical embodiment.

1.3 The Orthographic Conception. Now, what are these words that we transmit by means of their utterances and inscriptions? Well, the prevailing view, especially among those trained in the traditions of logic (as I am), is that words are the *types* of which utterances and inscriptions are *tokens*. This, I now think, is quite wrong. And misleading, even as a model.

It seems to me in many ways that this is a sort of updated

[8] I used to think that inscriptions, sometimes called *tokens* by logicians, were physical objects, possibly discontinuous, of the kinds: piles of ink, neon tubes, pieces of metal, and so on. Here is an example of such a discontinuous physical object.

NO

Then I realized that there is another way in which we can form an inscription of that word. You do it like this.

NO

Now here again I make use of another discontinuous physical object (discontinuous because of the piece remaining from the middle of the "O") in order to produce a second inscription of the word "NO", but notice that the inscription is not the discontinuous physical object which constitutes the stencil, it's the *space*.

I first realized this when I thought about inscribing words in stone, and I realized that you are creating a physical object when you inscribe a word in stone, but the token of the word is not the great big heavy physical object, the physical object which is the token of the word is the light-weight space. This must be the notion of space in the sense in which architects and sculptors create space by enclosing them with larger physical objects. Note that the enclosure doesn't have to be square for the space to be square. You can have a square space inside a sphere. And notice also that this is the kind of space that can be moved around, from place to place. The space is intimately connected with its enclosure but not identical with it. You could knock the corners off the enclosure without affecting the space.

version of the Platonic notion of abstract forms. The eternal, unchanging Platonic forms (shapes, perhaps) are the types, and their physical embodiments, which *reflect* these abstract forms, are the tokens. I think that the token/type model is the wrong model for the occurrence/word distinction (i.e. the utterance/word distinction or the inscription/word distinction). The token/type model best fits what I call the *orthographic* conception of a word, the typesetter's conception.

According to this conception, expressions of the language consist of strings of atoms called 'letters', certain strings form words. The letters are abstract entities whose tokens, for the typesetter, are individual pieces of type. (It's strange that on this conception, type is a token, but that does seem to be the result.) You surely know this conception very well. We were all taught it when we studied the syntax of formal languages (not to mention when we took print shop in high school). It belongs to the formal subject, formal syntax. And its study is the study of an algebra.

1.4 The Common Currency Conception. The Token/Type Model versus the Stage/Continuant Model. I propose a quite different model according to which utterances and inscriptions are *stages* of words, which are the *continuants* made up of these interpersonal stages along with some more mysterious *intra*personal stages. I want us to give up the token/type model in favor of a stage/continuant model. This is not, I think, simply another way of doing the metaphysics of types under the old token/type conception, but a quite different conception of the fundamental elements of language. I think of my conception of a word as a *naturalistic* conception. Because the interpersonal transmission of words is so central to my conception, I adopt a phrase of Kripke's, and I call my notion the *Common Currency* conception of a word.

Here's just one way in whch my conception differs from the token/type conception. On my conception, there is a single word "color" spelled one way in Canada: "c", "o", "l", "o", "u", "r" and another way in the United States, "c", "o", "l", "o", "r". Similarly, there is a single word which is pronounced *shedge-yule* in Canada, and *skedge-oo-ul* in the United States. (I believe that the English have also adopted the Canadian

pronunciation.) We are, of course, familiar with dialectal variation in pronunciation, but I call to your attention that there is also dialectal variation in spelling. This dialectal variation in spelling of the word "color" for example does not, repeat *not*, make for different words.

1.5 Dialect and Idiolect in Saying and Writing. I should point out that there are also idiolectal variations in pronunciation, as for example in the case of a speech defect. Some of us also have idiolectal variations in spelling. Society treats these two kinds of deviations—in speaking and in writing—in extremely different ways. We are very liberal nowadays about variations in pronunciation, at least in the United States. In England, although there are many and more disparate dialectal variations in speech, there seems more acceptance of the idea that there is a correct way to pronounce things. In the United States, even the national networks have anchormen whose pronunciation is quite different. Dan Rather of CBS speaks with a noticeable Texas, or Southwestern, accent—a regional dialect. Tom Brokow of NBC says *mi-we-un* where I would say *million*—an uncommon, but not otherwise unheard of, idiolect. And Peter Jennings of ABC says *about* for "about"—a Canadian dialect, I believe. These variants are tolerated.

We do not tolerate, however, idiolectal variations in spelling. Now it's been suggested to me that the variations in pronunciation we tolerate is regional, or group, variance—i.e. dialectal variance—and if there were a regional or group variance in spelling, that would be tolerated also. I don't believe that for a single minute. We tolerate Brokow's idiolect and all sorts of other strange and unique speech styles. I'm quite confident that children from the South, as they start learning to read and write, have a strong tendency to spell the two words "you all" "y", "a", "l", "l" (that's simply making a natural transcription). But that this charming spelling dialect is quickly suppressed by the forces of spelling bigotry and intolerance.

It must be obvious to you that I'm a person who has suffered—and who continues to suffer—serious discrimination regarding my idiolectal spelling. And I wish to point out that in, say, the seventeenth century, people were much more liberal about such things than we are today.

1.6 How words change: Platonic Shadows versus Evolving Nature. I make these points to dispel the lurking thought that a word has some fixed and perfect Platonic form, represented say by the way it is spelled, which is then embodied in the imperfect, variable, and changing ways in which people pronounce it.[9] Historically of course the matter was quite the opposite. Speech is prior and writing is a transcription of speech. So there is no metaphysically fixed form in either speech or spelling, no matter what social programs are mounted for standardization. There are spelling variations, there are pronunciation variations, there are all kinds of variations that take place over time.

On the other hand, on my conception, there are two phonographic words "base" (meaning '*low*' and *bottom*), not, as the orthographic conception would have it, a single word with two meanings. (And, as we will see, I do not simply define *word* as an orthographic word combined with a meaning.)

1.7 Creationist and Developmental Linguistics. I will want, eventually (but not in the present paper), to connect these speculations with the questions of whether the Babylonians believed that Hesperus is Phosphorus and whether Peter believes that Paderewski is not Paderewski. So let us consider the name "Hesperus". I have a story about how the word "Hesperus" came to us. The story is surely incorrect (after all, the Babylonians didn't speak Greek), but it gives the flavor of my views. I imagine that at some point some Babylonian looked up in the sky one evening and said (in Babylonian) "Oh, there's a beauty. Let's call it—", and then he introduced the name. What he did was to create a word. He created a word as a name, a tag, for this beautiful heavenly body. He then passed that word on to other people through inscriptions and utterances. Those people passed it on to others and so on (I'm going to talk more about this passage). As it went through different communities, the way this word was pro-

[9] I recently read a book about Darwin in which it was said that an enormous amount of evidence for evolutionary ideas was already available when Darwin was working. One of the things that stopped people from putting it all together was this Platonic idea of the fixed, eternal, and unchanging form setting the limits of variation for the shadowy objects of the sensible world. They were carried away from the thought that there could be natural kinds which went through the dramatic transmutations claimed by evolutionary theory. As I said earlier, I think of my conception as being naturalistic, as owing more to the theory of evolution than to algebra.

nounced and written changed in very dramatic ways, through whatever processes account for dialectal variation.

The presupposition of these processes of change are the principles of continuity in accordance with which a changing word retains its identity. As *we* pass through various communities at various stages in our lives, we also change dramatically. (I think I am probably more different now from the way I was when I was, say, eight yours old, than the word "Hesperus" is now from the way it was when it was, say, five minutes old.) But we still have the notion that we are a single entity. And so it is for the word. Changes in pronunciation and spelling need not suggest the notion of *replacement* of one word by another, which then takes up the task in the manner of a relay race. Rather, we can use the notion of a single entity undergoing *change*.

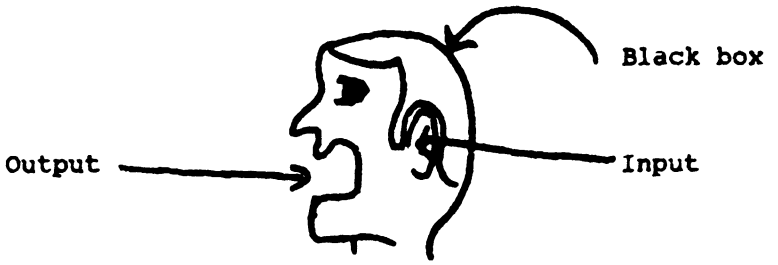
1.8 Intrapersonal Continuity. Now, I've talked about the process by which the word was transmitted *inter*personally. What about the process by which the word is transmitted *intra*personally? This is the most difficult stage to understand, at least for me to understand, because it is so deeply implicated in cognitive psychology. Suppose a particular word is transmitted from you to me. Now at some point I make a transmission to someone else. Question: Am I transmitting that very word?

Some word was transmitted to me by way of utterance or inscription. I transmit some word by way of utterance or inscription. We can phrase the question in this way: 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? We can thus turn the question into a question about the relationship between input and output utterances or inscriptions.

I don't mind if you want to continue to call utterances and inscriptions "tokens", although I'd prefer "utterance" or "inscription", so long as we do not get caught up in the metaphysics of the token/type model. Because it is beyond doubt that the utterance or inscription transmitted *could* be an utterance or inscription of the same word as that received, although the difference in phonographology, the difference in sound or shape or spelling, can be just about as great as you would like it to be.

"Just about as great as you would like it to be", that's a strong statement. I will try to defend it.

Let me restate the problem using the technical apparatus and terminology of cognitive science.



Something goes on between the reception of an utterance as an input and the transmission of a distinct utterance as an output. What happens in the black box during this intrapersonal processing, what is it that connects particular input and output circuits? What is it that makes a particular output, the transmission of the same word as that carried by a particular earlier input? I can't provide a detailed answer to this question, but I can give you some examples to show you how great the differences in sound, shape, or spelling can be in cases in which, unless we are completely dominated by the token/type model, we would agree that the word being transmitted is the same word as that received.

Consider this thought experiment: I say the name of an individual, possibly a name known to the person to whom I am speaking. The subject is to wait for a count of five, and then repeat the name. I say a name, then the subject says the name. I say the next name, then the subject says the next name. So, if I say "Rudolf", the person says "Rudolf"; "Alonzo"—"Alonzo"; "Bertrand"—"Bertrand", and so on. Because we have to worry that the subject might be, in Kripke's term, reticent, if he succeeds in repeating the name, we reward him with a dollar, or, if he has tenure already, a thousand dollars, enough at any rate to motivate him. I think that if we set up the story in this way—the subject is highly motivated, he is sincere, he is not reticent, he is

reflective—whatever that means, then we are very strongly inclined to say that when this person speaks, he is repeating the very name that he heard. I'm not saying what's happening inside the black box, I'm not saying *how* he does it, I'm just saying that from the description of the case it's clear that we would agree to describe his output as a *repetition* of *that* name. This notion of repetition is central to my conception.

There is a physical transmission of my output utterance, which is the subject's input utterance, to the subject's black box, then there is a psycho-physical (or better, a physico-psychological) transition where those sound waves hit the ear and something goes up and is put into what we call memory, and then, after five seconds, *it* (that something) is called out of memory, goes through a psycho-physical transition in the vicinity of the mouth and throat and the output utterance appears. (I am purposefully vague as to whether the *psycho* part takes place in the mind or in the nervous system, though I lean to the answer "both".)

The two psycho-physical transitions, at the ear and at the mouth, significantly affect the subject's ability to match the physical properties of the sound waves that go out (his output) to the physical properties of the sound waves that come in (his input). (I am assuming here that sound waves with 'matched' physical properties will 'sound' the same to each single auditor.) For example, on the output side, I may say the names in my deep and profound voice, and the subject may not have the apparatus to produce a deep and profound voice. He has a high squeaky voice and that's the best that he can do. If you happen to know someone, as I do, who has a certain kind of congenital hearing defect, you know that although his or her productive mechanism may be perfectly normal, the problem at the input side produces noticeable and characteristic kinds of changes from the sound pattern that come in to the sound pattern that goes out. (At least, as I hear those sound patterns.)

So individual differences in the physiological processing at these psycho-physical transition points may make it that what comes out is not going to *resemble* what went in. However, and this is what is important for this thought experiment, the exact functioning of the psycho-physical transition mechanisms is irrelevant to our characterization, our intentional characterization, of the case as one of '*repeating* the word he heard'. He

may not do it well, from the external point of view. But it *is* what he is doing. We don't question that the connection has been made in the black box. No matter how poor the subject's *imitative* ability (his ability to make his output resemble his input), we can imagine circumstances in which we would say, "Yes, he is repeating that name; he is saying it in the best way that he can."

Contrast this with a wealthy mischievous subject who has decided that he's going to play a trick on us and instead of repeating the names as he hears them, ignores the input and just utters names at random (or, he may have prepared his own list ahead of time which he recites in order). Even if, by happenstance, the sounds that come out in these two cases equally resemble the sounds that went in, the first case is a case of repetition and the second is not.

The identification of a word uttered or inscribed with one heard or read is not a matter of resemblance between the two physical embodiments (the two utterances, the two inscriptions, or the one utterance and one inscription[10]). Rather it is a matter of intrapersonal continuity, a matter of intention: Was it *repetition*? We depend heavily on resemblance between utterances and inscriptions [using *resemblance* here not to mean matching of physical characteristics but of their appearance as we look and listen] in order to divine these critical intentions. If it sounds like "duck", it probably is "duck". But we also take account of accent and idiolect and all the usual clues to intention. It is the latter that decides the matter.

In fact, as Philip Bricker has pointed out, when we *repeat* what someone has said, we don't aim to imitate the pronunciation, we aim to standardize it (by our own standards). Imagine asking a third party to repeat what a speaker with an unintelligible accent has said. Would he *imitate* the speaker?

There is a story about a missionary who mispronounced certain verbs of the local language, with the result that his sermon sounded like an exhortation to highly questionable behavior. All this, much to the amusement of the natives. But what was the missionary *saying*? Is what we say really so subject

[10] One of the events that set me thinking along these lines was a talk by Arthur Danto in which he asked, as I recall it: In what way do an utterance and an inscription of a single word resemble one another?

to the vicissitudes of ear and tongue, or to the dialect of local custom? I would claim that a *mis*pronunciation of a word is an utterance (a *pronunciation*, if you will) of *that* word. It seems to me that cases in which one utters a word other than the one intended must be exotic indeed.[11]

There are, of course, familiar ways in which one can fail to say the word one intends to say, for example by intending to say the magic word that opens the cave door, but forgetting exactly which word that is. Or intending to say the name of an old acquaintance at a party, but drawing a blank. (This last is a looming fact of life for many of us.) Here the word is given under a concept, as Frege would say. This is quite a different thing from mispronunciation, even startling, inexplicable mispronunciation.

In view of the fact that individual differences in physiological processing at the psycho-physical transition points may affect the *resemblance* of output to input, we can imagine somehow getting into the transition mechanisms of the subject and putting filters of various kinds on them so that we get—and this was the claim that I made earlier—we get differences in sound *just about as great as we would like* between what comes in and what goes out. If, however, we were convinced that the source of the change was as described—due to the filters at the psycho-physical transition points—and that the cognitive link that has to take place inside the black box was in order, we would still say it's the same word. It's just that the subject can't come anywhere close to giving the word a standard pronunciation.[12]

[11] The exotic cases I have in mind are those in which *to the astonishment of the speaker* the wrong word came out. I have been witness to, and subject to, this experience on several occasions in connections with proper names. "Wait a minute", the speaker says, "Did I just say 'Eleanor'? I meant 'Harriett'." Some dark force has reached into the speaker's psyche and misdirected the hand of intention.

[12] We are familiar with everyday occurrences of this in connection with people who have different native languages and who have a different repertoire of sounds. One of my Japanese friends, who spoke unaccented Californian, was trying to explain to me how to say two of my favourite words, one of which is "Netsuke" and the other is "Hokusai". There is a "u", as we write it in English, in both of those words which doesn't exactly disappear, and isn't exactly sounded. He kept saying "You are saying this [and he would imitate my pronunciation]; you should be saying this [and he would pronounce the word 'correctly']". I couldn't hear the difference between his imitation and the 'correct' pronunciation. Conversely, as we know, some of our Japanese friends have great difficulty with the R-L distinction, a distinction that we easily make.

1.9 The Criterion of Word Identity is not *Resemblance.* My chief point is this. There is a word taken in; there is a word sent out. There is a mystery about what has to go on inside the black box in order for it to be the same word, a mystery about how to analyze in detail what must go on in order for it to be the same word. (I'm imagining an analysis in the language of psychology, not brain physiology; an analysis in programming, software terms, not in electronic circuitry, hardware terms.) But the first thing that we should get out of *our* heads is the idea that we can tell whether the input and the output are utterances of the same word by looking at (or listening to) the physical object that comes out, and looking at (or listening to) the physical object that goes in, and trying to make a phonographic comparison of the two to see whether they are similar enough in some specified way.

1.10 Continuity through Memory is Hard to Trace. When the word goes through the black box, when the word is received from one person and stored for passage on to the next person, it isn't, of course, put into the pocket in the way in which a coin can be stored in its passage from person to person. The coin is put into the pocket and there it is located. There is a definite answer, whether we know what it is or not, as to whether the lucky coin that your coach gave you is really the very one that his coach gave him, or whether it is a different coin that looks pretty much like the lucky coin (in fact, is indiscernible from it, in the way of modern artifacts).

In the case of the word, we feel that the comparable question doesn't have the same very straightforward answer, because it isn't put into the pocket, it is put into memory. Remember, in our experiment we said to the subject "I'm going to say a name. Then *wait for a count of five* and repeat the name". This form of storage, *in the mind* (rather than in the pocket), makes the continuity much harder to trace.

1.11 Short Circuits and Gaps. On my theory of words there are two special problems having to do with the way in which the words are stored in memory that make for particular difficulties. Each problem corresponds to a kind of error regarding word individuation to which even a competent language user is susceptible. The first kind of error is this. A transmission of a

word may come into the black box and then another transmission of the same word may come into the black box, and inside the black box this very same word might get stored in two different locations, *stored*, so to speak, as the transmissions of two different words. If I say "tomato" and you say "tomahto", a naive subject may take us to be speaking of different vegetables. It really is just one word. There have been two receptions of the same word, but it is not stored as a single word. It is stored as two different words.

Somehow, in the black box, the different branches of the same word (i.e. the different input utterances of the same word) weren't all properly linked together. And that means that when the black box emits, it thinks it has two words, and it will make a choice as to *which* 'word' it is going to emit.

The converse phenomenon also occurs. Consider a speaker with three friends, "Mary", "Merry", and "Marry" as he calls them (the last might be a nickname for Zsa Zsa Gabor). Now in some American dialects these three words are pronounced indistinguishably; they are homophones. From a psychological point of view, from the point of view, as it were, of the black box of the speaker, three quite different words are being uttered. He may even think he is pronouncing them quite differently, he has different spelling for them, he puts his tongue in a different place when he speaks. But from the point of view of the listener, what comes out is *at best* three homophones.

To see how easily this can occur, consider a case in which no error is made. Let's take an example of someone who has two friends, two distinct friends, named "John" and "Jon". Then the person can choose to say "Look, it's tautological that John is no taller than John. And also that Jon is no taller than Jon. But it is not the case that Jon is no taller than John." What has he done? He has these two names, they're homophones, and he makes two choices of the first name and produces a sentence of the form "*a R a*". He then makes two choices of the second one and produces a sentence of the form, shall we say, "*b R b*". And then he makes one choice of the one and one choice of the other, and he produces a sentence of the form "*a R b*", in terms of the very words that were used. (Suppose the relation *R* were *identity*.) It is important to recognize that having achieved this insight, we can kick away the spelling ladder.

This situation *can* arise. If we do kick away the spelling ladder and spell both names the same way, we have two, common currency, phonographic names for two different people. One of these names was created (i.e. introduced as *that* name, a name of that person), say, by the first friend's mother, and the other one was created, say, by the second friend's grandfather. The two names have had rather different life histories, but they met when both were stored in the black box of the common friend. Within the black box they were correctly stored as being two different words, and then when the black box emits, it chooses which word it is going to emit, which 'circuit' it is going to continue.[13]

But note that the very same phenomenon could easily occur when there was only one name and the person made an error in thinking that there were two different names, in thinking that there were two different common currency names. Imagine now that there was only one common currency word, but that it was stored in this way as if it were two. An error in word individuation is being made by the person whom this black box represents, an error, as I would call it, in syntactical-lexical form. He is, of course, a perfectly competent speaker of the language, a native speaker in fact. This error that he is making is not really to be held against him, because it could happen to any of us. There are so many people to be named and so few *generic* names to go around. (Generic names are the genera, or species, of our individual common currency names. More about this later.) Errors of the first kind are taking one name to be two.

Peter is making this kind of error in Saul Kripke's "Paderewski" case, mentioned in the preface. It is my belief that the analysis in terms of word individuation is valuable, and perhaps critical, in understanding that fascinating case and in distinguishing it from the more familiar Hesperus-Phosphorus cases in which there is no problem of word individuation. I don't claim that all mistaken identity cases,-not even all that pose

[13] Again, an actual case: My mother's primary care physician is Dr. Shapiro. He referred her to a specialist, another 'Dr. Shapiro' as it happened. My mother reported her gratitude to Dr. Shapiro for sending her to Dr. Shapiro and compared Dr. Shapiro's virtues to those of Dr. Shapiro in a blithe piece of discourse, clearly oblivious to the homonomy. I was racing to keep up (which I was strangely able to do). But from her point of view, she was quite properly using two different words to refer to two different people. Why should there be a problem?

problems for direct reference theory, resolve to mistakes about the identity of words. But this one does.[14,15]

There is a second kind of error to which we are all susceptible. Errors of the second kind occur when there actually are two persons with phonographic, common currency names (but still distinct names, one created by the mother and the other by the grandfather), and the two common currency names are linked and stored in a single location (and it is from that single location that all future utterances of "John" will arise). This is an error of the second kind: taking two names to be one.

The second kind of error is a short circuit—two different circuits got wired together—zap, the whole thing goes up in smoke. I am inclined to think that when two different common currency words are wired together in this way in a given black box, which then pulls from that common source and transmits, nothing whatsoever is being said. Is it transmitting the first word? Is it transmitting the second word? I think there is just no answer to that question. The two words have been co-mingled in such a way that there is just no answer. Harking back to earlier remarks about the difference between confusing the identity of persons and confusing the identity of words, I note that I am not claiming that in errors of the second kind we cannot tell what the speaker means, or to whom the speaker is referring. It is rather that, even if we could identify the referent, we could not thereby infer the identity of the word.[16]

[14] How are we to think about Kripke's famous London-*Londre* case? Should we regard it as a case of dialectal variance, thus assimilating it to the Paderewski case? Or should we regard "London" and "*Londre*" as distinct words, and treat it as a Hesperus-Phosphorus case (or perhaps a Germany-*Deutschland* case)?

[15] A confused identity case that does *not* seem to resolve to a mistake about the identity of words is that described by Russell in 'On Denoting' to interpret the results of giving "the author of Waverly" *primary* scope in "George IV wished to know whether Scott was the author of Waverly" (a case of quantification in). "This would be true" Russell says "if George IV had seen Scott at a distance, and had asked, 'Is that Scott' ". George IV's innocent confusion, expressed in "Is that Scott?" seems to me to involve no *syntactical-lexical* confusions, as I would put it. At the moment, the case also seems to me to pose no serious threat to direct reference theory. (Oddly enough, it does pose a serious threat to Russell's own theory, since the primary scope interpretation allows us to prove the sentence attributing an interest in the law of identity to the first gentleman of Europe. Try it, using "Scott" as a genuine name.)

[16] Keith Donnellan in section ix of 'Proper Names and Identifying Descriptions' (*Synthese* 21 (1970): 335–58, reprinted in D. Davidson and G. Harman (editors) *Semantics of Natural Language* (Humanities Press, 1972)) gives us the mistaken identity case of J. L. Aston-Martin. A party-goer mistakenly takes another man at the party to be the

Errors of this kind are quite different from Paderewski cases of the first kind in which there is a single word that is being transmitted but the speaker makes the mistake of thinking it to be two words.

II

WHAT ARE NAMES?

2.1 Names as Words. I have spoken of *words*, though my examples have often involved *names*. And truth to tell, it is *names* at which I aim. It is names that have been thought to challenge direct reference theory.

Names are a special kind of word, so special that some have thought them not to be a part of a language at all. I disagree with this and will emphasize ways in which names are like other words, but I do not disagree that names are special in several ways. Even if one were to conclude that names are so unlike other words as not to be regarded as a part of any particular language, this should not count against applying the earlier principles of individuation to names.

All that I have said of the interpersonal transmission and intrapersonal processing of common currency words seems to me to carry over directly to names. Remember Paderewski, Jon and John, and my mother's two Drs. Shapiro.

Furthermore, it seems clear to me that my processing of my friends' names is *correct.* They *are* different common currency words. Different people created their names and did it on different occasions, so it seems fitting to say that they have different names, that their names are different common currency words.

Once, when I gave a talk on these matters, Paul Benacerraf said to me, in connection with David Israel who was in the audience at the time, "I thought that you and David had the same name, now you tell me you have different names". John

famous philosopher J. L. Aston-Martin. In the party-goer's subsequent discourse, Donnellan identifies the referents of successive uses of the name as variously: the man at the party or the famous philosopher. His technique is to use the indices of intention to identify the topic of discourse. No sorting of words occurs, as is obvious from the fact that in the case, as described, only a single name is involved. It is interesting to note that Donnellan's techniques would seemingly work just as well in identifying distinct referents even if one of the two referring expressions were anaphoric on the other.

Perry put it somewhat more strongly afterwards saying, "Your views are insane because there's just this one name 'David', which has been around since biblical time".

It might be thought that this observation counts against my conception of words. Not at all. These ancient names to which Perry alludes are naturalistic objects. They live in the world, not in Plato's heaven. They are cultural artifacts, created by us, transmitted by us, stored by us. Surely, there is no argument here for the token/type model.

At most what we have here is a disagreement about subtleties of individuation. And perhaps not so serious a disagreement as to be invulnerable to a conceptual distinction. Let us acknowledge this other conception of a name, which I call a *generic name*, just as we acknowledged the orthographic conception of a name. The notion of a generic name is useful for clearing up the apparent disagreement over when two utterances are utterances of the *same* name, and it may be useful for some practical purposes—for spelling checkers and typographer, things like that. But for serious semantics, I think that it is my common currency conception that would be important.

There is the generic name "David", and then there is my name "David", there is David Lewis' name "David", and so on. These three are all distinct words. The latter two have—and here I speak carefully—a semantic function: They *name* someone. The first, the generic name doesn't name anyone (doesn't name any*one*, perhaps it names or is an unnatural kind). Furthermore, it doesn't pretend to name anyone (as certain empty common currency names do).

Generic names may be the closest thing in my theory to the Platonic forms that word types were said to be by the token/type theorists. (I personally prefer the species/specimen analogy, and don't forget the evolution of species.) But even if you wish to think of generic names as types and my common currency names as their 'tokens' (ugh!), two utterances of *my* name "David" are utterances of the *same* 'token' of the generic name, whereas an utterance of my name and an utterance of David Israel's name are utterances of *different* 'tokens'. To put the point another way, no matter how you slice it, the individuation of common currency names, as I have described it, must be taken into account.

Thus, I encourage the idea that in addition to the common currency names—these distinct, fairly recent names, David Israel's and mine for example, which were created at different points, and which have had distinct life histories—in addition to the common currency names, there is also another *kind* of word which we call a 'generic name'.

2.2 Causal Chains and Generic Names. The recognition of generic names poses a delicate problem. When people think about the picture of the transmission of words that I have tried to outline here, they call it the *causal chain* theory, or something like that, and they think that the question of intrapersonal continuity is just a question of whether the utterances of the name that was emitted was caused by the utterance of a name that was received. But you see, there are very delicate issues as to what goes on in the black box and the nature of the causation. Indeed, when I was named "David", my parents had David Hume in mind, and they so admired him that they thought, "Let's name our son with the same name". So I was named *in honor of* David Hume. My parents didn't just make up the name "David", as if they were sitting there trying to think of a name for the baby, and they suddenly said, "Duh, Duh, Day, Day, Dave, Dave, David, That's it! David!". That is not what happened. There was a pre-existing generic name "David". My parents were aware of it and of many of its associations, including the fact that the common currency name of the great philosopher David Hume was to be drawn from it. That is how they thought about the generic name. Having David Hume's common currency name in mind, and *in honor of* its referent, they decided, as is our custom, to name their child with a common currency name drawn from the same generic name. That is the sense in which there is a causal connection between my common currency name and that of David Hume. But such a connection is clearly not of the right sort for the names to be the same. The earlier discussion of the notion of *repetition* was an effort to show that the so-called causal link fell in the realm of the intentional, and to discriminate it from certain other intentional connections like being named *in honor of* another.

2.3 Are Names Parts of a Language? When I first started thinking about generic names I was aware of a lot of literature about

names that says that names are not part of any language because one can just create new names at will, whereas the lexicon of words other than names forms a relatively stable body characteristic of the language. If you think in a careful way about common currency names and generic names, you will recognize that this is not correct. Common currency names can be created at will. But so far as I know, natural languages have, at a given time, a fairly fixed stock in its lexicon of generic names.[17] When you expect a child you go to the bookstore and you buy a book called something like "What to Name the Baby; 3,000 Generic Names of English". Just as the dictionary lists and 'defines' 50,000 *words* of English, the generic names book will list generic names and provide some lore about each name, such as information about several famous people whose common currency names have been drawn from it. When we name people, we usually draw from this relatively small, finite lexicon of generic names.

The idea that you can freely name someone anything you choose is very misleading. "What shall we name the baby, dear?" "How about Tkbtkbtkbt?" No way! It wouldn't be English. How would you pronounce it? "Tkbtkbtkbt" just isn't a generic name of English. You can't just decide you are going to name a baby with that. There are a lot of names, you have a lot of choices, but you can't name it that.

My point here is that names, like other words, must subscribe to certain regularities. Certain of these regularities have to do with admissible sound and spelling patterns, and others are simple matters of social control. The free creation of names is possible only to the degree that the linguistic community will tolerate it. And the degree of tolerance in a linguistic community for linguistic deviance in naming practice is an *empirical* question, not one to be solved by philosophical analysis.

> PARIS—Marc Borneck expected some interesting reactions when he named his daughter Prune last fall. He hasn't been disappointed.
>
> To begin with, his mother-in-law hated the name

[17] I am thinking here of generic names for persons. There are also 'rules' constraining the generic names for racing horses, show dogs, etc. John M. Carroll's *What's in a Name* (W. H. Freeman, 1985) contains much interesting information about naming practices.

(which in French means plum, not prune). But little Prune turned out to be so winning that Grandmother soon was won over to the name, too.

The neighbors were taken aback. But Mr. Borneck is a beekeeper by profession, and they figure—rightly—that this was just another of his efforts to get his children to love nature.

But the people who were really outraged were in the French government. And that was what mattered the most.

The local prosecuting attorney informed the Bornecks that the name was not just absurd but also illegal. He sent policemen to the Borneck home to confiscate their family book, a quasi-official register of important family events, in which local registrars had inscribed the name Prune. And he ordered that the girl henceforth be known by her middle names, Mae Kim.

He didn't like those much, either, and says he could have voided them, too—substituting, if necessary, names of his own choosing. But he says he wanted to be a nice guy about it.

The Bornecks, who still call their daughter Prune, have taken the matter to court, and so far they have lost. They have discovered, like other unconventional parents before them, that first names in France are strictly regulated by law—a law drafted in 1803 under Napoleon, who disliked offbeat names although his own didn't seem to hold him back much.

"Personally, I think 'Prune' is kind of cute," says Michele Signoret, A Justice Ministry official. "But is it a name? That's up to the courts."[18]

2.4 Promises Kept by the Mischevous Babylonian. Following the discussion of generic names, I am now in a position to fulfill my promise to show how there can be distinct names which are phonographs and which also have the same semantic value. Let

[18] *The Wall Street Journal*, Western Edition, June 10, 1987. I am informed that in Mussolini's Italy a law was passed forbidding the naming of children with any (generic) name that had already been used. This contrasts with the Netherlands in which you are not permitted to use any name that has *never* been used. (Perhaps the Italian law only forbade naming children after the parents.)

me tell you about the case of the mischievous Babylonian. One evening, the mischievous Babylonian looked up and saw Venus, and he thought to himself "This one is just as beautiful as Phosphorus, so let's call it 'Phosphorus' too". Now maybe he actually knew he was naming the same thing; maybe he was a time traveller playing a little joke on his Babylonian friends. But more likely, he didn't know. He just decided that he would name this rather attractive 'star' *in honor of* the great and beautiful heavenly body seen in the morning in a rather different location during a different season. So he names, or perhaps we should say *renames*, Venus "Phosphorus". We may suppose that this name passes muster with the Babylonian Justice Ministry and comes into common use. Now it seems clear that we have two common currency names "Phosphorus", one somewhat older than the other, and that they start out, at least, as phonographs. Who knows, after a little while they may drift apart in terms of pronunciation because as astronomers talk more and more about the sky, they might feel that it is confusing to have the same generic name "Phosphorus" for two 'different' heavenly bodies. So they might add suffixes and start calling them "Phosphorus I" and "Phosphorus II". But for a little while, they were distinct phonographic common currency names with the same semantic value.[19]

2.5 Are Common Currency Names Just Generic Names + Referents? I want explicitly to disavow one possible misapprehension about my notion of a common currency word. It is *not* simply the notion of an orthographic word combined with a meaning, as some have thought because of the way I distinguish common currency names from generic names. They have thought that a generic name was just an orthographic word, and a common currency name was just an orthographic combined with a meaning (or a referent, insofar as they differ). No!

As noted above, generic names are also natural objects. They undergo the same changes in spelling and pronunciation that the individual common currency names do, more so because

[19] I have recently become aware of a similar claim in footnote 28 of Saul Kripke's *Naming and Necessity* first published in G. Harman and D. Davidson (eds) *Semantics of Natural Language* (Dordrecht: Reidel, 1972); revised edition published as a separate monograph, *Naming and Necessity* (Oxford: Basil Blackwell, 1980). Reference is to the revised edition.

they are older and range more widely. Generic names are not orthographic words. In fact, they almost are the common currency names but for a difference in individuation. I say *almost* because there are inscriptions of generic names that are not inscriptions of any common currency name (as in *What to Name the Baby*). These matters are subtle (and perhaps are slipping out of my grasp).[20]

This suggests another incorrect hypothesis: that common currency names are generic names combined with a meaning (or a referent). This is more plausible because generic names have the appropriate wordly character but are entirely without meaning (they are empty vessels). However, this analysis would not give us the *actual* common currency names of a natural language. On this conception there is a *possible* name which combines me, as referent, with the generic name "Phosphorus". But I have never been named that.

So suppose we try, as a last ditch attempt: A common currency name is a generic name combined with a referent that has actually been *given* that name. (Don't come down too heavily on 'given name', that's another topic.) But even this is wrong. The case of the mischievous Babylonian makes it clear that *distinct* common currency names can share the *same* generic name (i.e. be of the same genus) and have the *same* referent.

All these proposals to understand common currency names as 'combinations' miss an essential feature of my conception. Common currency names (and other common currency words) are not abstract constructions, they are natural objects. Not physical objects, though most will have physical embodiment at many places and times. And not mental objects, though most will have mental embodiment (an oxymoron?) at many places and times.

[20] Suppose the common currency name used to dub the baby is not drawn from any pre-existing generic name. Then a new generic name is introduced in the act of dubbing the baby with its common currency name. Now suppose that the common currency name flourishes, but that the generic name is never used (i.e. never uttered or inscribed) except through uses of the common currency name. What exactly does the difference between the two names consist in? They must be distinct since the generic name doesn't name anyone and also has a capacity (even if unexercised, *ex hypothesi*)—to 'generate' other common currency names—that the given common currency name lacks. This problem seems analogous to asking about the ontological status of a new species of animal of which only one specimen ever exists.

One might think of them as trees. Stemming out from their creation, with physical and mental segments; the mental segments able to produce many physical branches and able to merge many physical branches, and the physical segments each stemming from a unique mental segment and able to produce many mental branches. (Perhaps I need a diagram). At any rate, they are objects of the created realm, created by language makers. The world is not brimming with unspoken words. Words never actually created *are* not.

2.6 Can Common Currency Words Change in Meaning? Or, to be more specific and more directly relevant to the purposes I have in mind, can a common currency *name* undergo a change in referent? There is no *prima facie* reason against it. I re-emphasize: The identity of a common currency word lies in its continuity, both interpersonal and intrapersonal, as has been discussed. It is a matter for further analysis to say whether such an entity could change meaning (or reference). It is certainly no part of my conception that it cannot.

The matter does, however, call for careful thought. One might consider two kinds of polar cases: In one case you intend to use (to *repeat*) a given common currency name with whatever referent it may have. ("What is Hesperus?" you ask, overhearing a conversation in which the name is used.) In the antipodal case, you intend to dub a particular thing using an apt generic name. In the former case there is continuity, in the latter, creativity, a new name is created. But there are those troubling cases (first thrust upon our consciousness by Keith Donnellan, and then Gricefully reconceptualized by Saul Kripke) that seem to lie in between: the man with the Martini, the false introduction, and their ilk.

It would be gratifying to be able to show that the process whereby a common currency name appears to change its referent involves these middling and conflicting intentions in such a way that when we are prepared to say that utterances of a name now have a new referent, there will have been sufficient weight given to the creative side of the intentions for us to claim that a new name has replaced the old. It might thus turn out that names don't in fact change in referent.

So one strategy would be this. Try to show that something like Donnellan's *referential use* is involved in cases of apparent change of referent of a given common currency name. Then try to show that this sort of referential use is sufficiently like creating a new common currency name from the genus of a given name so that by the time a new 'semantic referent' appears, a new name does also. I don't know that such a strategy would be successful.

2.7 Essential Properties of Names. One consequence of my view of names as wordly creatures may be to renew interest in some matters that had seemed obvious or irrelevant.

In 'Uber Sinn und Bedeutung' Frege argues, as against his earlier 'Begriffschrift' view, that "Hesperus = Phosphorus" could not mean that the name "Hesperus" is co-referential with the name "Phosphorus" because the connection between a name and its referent is arbitrary. "You cannot forbid the use of an arbitrarily produced process or object as a sign for something else." Frege concludes that no "genuine" knowledge would then be expressed by "Hesperus = Phosphorus", only knowledge of what one might call an accident of human culture.

But is it a mere accident that our name "Hesperus" names the planet Venus? To my way of thinking that is like asking whether Da Vinci's painting of the Mona Lisa might have been a picture of a horse. Certainly Da Vinci might have painted a horse instead of a woman. As Frege might say, "You cannot forbid a painter from arbitrarily deciding what to paint." But would the resulting painting have been the Mona Lisa? I won't try to answer this question. But I note that though Venus might never have been named "Hesperus" (it might never have been named at all), and though Alpha Centauri might have been named "Hesperus" (it might even have been named "Hesperus" due to a last minute change of mind by the Da Vinci-like Babylonian who in fact named Venus "Hesperus"), it does not follow that *our* name "Hesperus" might have named Alpha Centauri, at least not on my view of names.[21]

The question, "Is it possible that a name which in fact names a given individual, might have named a different individual?"

[21] One refinement: there may be reason to think that our name "Hesperus" could have *come* to name Alpha Centauri. My question is: At its creation, could it have named Alpha Centauri?

is, for me, a substantial metaphysical question about the essence of a common currency name. By calling the question a *substantial*, *metaphysical* one, I do not intend to puff up its importance nor to make it seem mysterious or occult. Perhaps, in the end, the question calls only for a decision, or perhaps, in the end, the question will seem unimportant. This may be the *tao* of substantial metaphysical questions. But there is not, I believe, an *obviously* correct answer.[22]

[22] The last minute assistance of Joseph Almog and Keith Donnellan has done much to make this readable.