

## Chapter Twelve Rejecting Identity Judgments and Fregean Modes<sup>1</sup>

### ' 12.1 Introduction

I would like to understand what the basic principles are that distinguish the vision of thought we have generated using the Strawson image of sameness marking from Frege's original vision of thoughts as exemplifying modes of presentation. The first conclusion I will reach is that, surprisingly, the way the Strawson markers mark identity plays no role in determining this difference. The interaction of Strawson's image of sameness marking with Frege's vision of modes of presentation yields strikingly unFregean results. Yet these results are not merely an artifact of the Strawson model. They follow given any model of sameness marking. Strawson's way of marking identity highlights a general feature implicitly present in all other models as well. It will take a while to argue for this conclusion. I will place particular emphasis on the equals marker, and on the image of thoughts as sentence-like, in which the equals marker model is embedded. For initially it is quite unintuitive that this particular model is isomorphic to the Strawson model. Such is the hold that the mental sentence image of thought has on all of us, with its careful but, as I will argue, illusory distinction between duplicates markers and equals markers, that is, between graspings of necessary identity and contingent judgments of identity.

Further search is thus needed to understand the division between the vision of thought we have generated and Frege's original vision of thoughts as exemplifying modes of presentation. What exactly is the source of the difficulties we have encountered in trying to interpret what a mode of presentation might actually be in a thinking mind or brain? I will argue that the classical notion of modes of presentation rests on two assumptions, both of which are mistaken. One classical source is an implicit denial that the way the mind uses the thoughts or ideas that it harbors has any bearing on their intentional contents. In particular, as suggested in ' 8.7, the Fregean model invokes the passive picture theory of the act of understanding sameness of content. What the mind does with the pictures is not involved in determining their contents, or in determining whether they are thought of as same or different in content. The second source is an internalist view of thought content, that is, a denial that the natural informational content carried by a thought has any bearing on its intentional content.

### ' 12.2 Does it Actually Matter How Sameness Is Marked?

Begin by considering duplicates markers. How will a system consisting, say, of mental sentences, and that uses only duplicates markers, come to realize that Cicero is Tully? It must put all the Cicero and Tully information into sentences using the same mental name, either "Cicero" or "Tully," choose which. Just as one of the dots has to go on the Strawson model, one of the mental words has to go on the duplicates model. So if it should turn out later that Cicero is not in fact Tully, whichever mental name got chosen will be equivocal, nor will the news that Cicero is not in fact Tully represent, for the

---

1 Parts of this chapter are revised from "Images of Identity" (Millikan 1997b) with the kind permission of Oxford University Press and from "On Mentalese Orthography" (Millikan 1993b) with the kind permission of Basil Blackwell Publishers.

system, any definite instructions for separating the information again into two piles. Duplicates markers do not differ from Strawson markers in function.

A moment's reflection shows that the synchrony model, the Christmas lights model and the anaphor model of sameness marking also have this result. Each merely binds all the things known about an object into one bundle. Each performs acts of identifying merely by merging bundles, so that no particular information about any subject has a particular or different mode of presentation from any other. The only method of marking we considered that does not obviously have this effect is the equals model of marking. We find this last model illustrated, implicitly, in Frege himself.

Frege might be interpreted as having supposed that the mind uses, in part, a duplicates system of sameness marking. For although senses were not supposed to be psychological entities, *graspings* of them surely are dated, psychological occurrences, and Frege seems to have held that it is awarenesses of duplicate graspings-of-sense that keep us from contradiction and govern the performance of rational mediate inferences (Chapter Nine). If so, this constitutes, I have argued, a substantial psychological claim. A perverse deity might have made our minds otherwise. Imagine, for example, the same sense coming into mental view twice simultaneously as subject term of contradictory judgments, but the demon has determined that only synchronously vibrating viewings of the same sense will move the mind to visage sameness of reference.

It is clear, however, that Frege did not view the result of an identity judgment to be the elimination from use of one of the two kinds of senses grasped, nor did he suppose any senses were equivocal. Rather, he took there to be two kinds of identity judgments, the "informative" ones such as "Cicero is Tully", and those such as "Tully is Tully" which are not. We might suspect, then, that it is the introduction of this new way of marking identity, used for identities not known *a priori*, that allows the Fregean thinker to identify referents without merging his thoughts of them together. This second identity marker, we suppose, functions like a mental equals sign. It marks two thoughts as being thoughts of the same, not by merging or destroying either, but simply by flagging them for use together in mediate inference.

The suggestion that there is something like an equals sign in thought which marks identity comes, more generally, from modeling thoughts on sentences, and it deserves very careful study. The sentence model is so deeply ingrained both in our everyday and our philosophical thought about thought that we will do well to understand it very explicitly. Otherwise it may mislead us in important ways. A way to begin to explore this model is to examine it in its most naked form, namely, that in which discursive thinking is analogized to the unfolding of a formal system.

### ' 12.3 Formal Systems as Models for Thought

A formal system is usually laid out in the following way. First you say what elementary symbols will be used: p's and q's, say, or A's and B's, and x's and y's, wedges, horseshoes, parentheses, and so forth. Next you explain how to construct well-formed formulas (WFFs) from these ingredients, typically with the aid of recursion. Then you may (but need not) lay down axioms. And last, you lay down rules (this in the

metalanguage, of course) that will move you from WFFs already laid down or derived to new WFFs. This laying down of symbols, well-formedness rules, axioms, and inference rules, is traditionally done by displaying tokens of symbols, using these as examples of the types of symbols to be put down and manipulated in accordance with the system's rules.

But how are other tokens of these same symbol types to be recognized? By what criterion will they be tokens of the same types? Typically, nothing is said about this. One supposes, traditionally, that some understood but unmentioned parameter on sameness of shape is what binds the tokens into types. For our purposes it will be important explicitly to recognize this implicitly designated part of a formal system. Formal systems have, I will say, besides (1) basic WFFs and rules for constructing more WFFs, (2) axioms or postulates and (3) rules of inference, also (4) "symbol-typing rules" or just "typing rules," that is, implicit rules telling what is to count as the same symbol or WFF again.

It is well known that in developing formal systems, rules can in general be substituted for axioms or postulates and also, axioms or postulates can be substituted for rules, though the latter (since the tortoise's historic conversation with AchillesCCarroll 1894), it is supposed, not without residue. Typing rules have generally been ignored. But in fact, as I will illustrate, typing rules also can sometimes replace axioms or inference rules.

A very simple interaction among axioms, inference rules and typing rules happens to be the interaction of the axiom, say, " $B=B$ " with the two inference rules "replace ... $B$ ... with ... $B$ ..." and "replace ... $B$ ... with ... $B$ ..." and with the typing rule " $B$  and  $B$  count as symbols of the same type." Now it is usually supposed that although one can trade off rules and axioms to a degree in this way, still the difference between the systems that result from these swaps is objective. Certainly, it is not so that there is really no difference between an axiom and a rule or no difference between a rule of inference and a typing rule. It is just that in some cases one can substitute one for another to determine the same set of theorems. What I will be urging, however, is that distinctions among these three categories break down in crucial ways when we hypothesize representation in the mind or brain. My conclusion will be that there is no difference in this context between a mental equals-sign marker, an identity rule and a duplicates marker.

To quell the suspicion that identity is somehow a special case here, let me first illustrate the trade off between typing rules and axioms and/or inference rules with a completely different kind of example. Consider, first, laws of commutativity. Standard renderings of the propositional calculus require that the equivalence of  $A\&B$  to  $B\&A$  and of  $AwB$  to  $BwA$  be either introduced as axioms, derived as theorems, or (aberrant but possible) given as special rules of inference. Suppose, however, that one were to construct a system in which the difference between left and right on the paper is ignored when grouping symbol tokens into types. No distinction is drawn between " $p$ " and " $q$ ", or between " $b$ " and " $d$ " and so forth. Similarly, " $AeH$ " is the same string as " $HdA$ ". More interesting, " $AwH$ " is the same string as " $HwA$ " and, if we use the traditional dot instead of

"&", "A@H" is the same string as "H@A". Here a symbol typing rule does duty for one or a couple of axioms, theorems, or rules.

For another example, suppose we read right-left distinctions as usual and play instead with up and down. We read "p" as a symbol of the same type as "b" except that it has been turned upside down. Then we use turning upside down for the negation transformation. We negate propositional constants and variables by turning them upside down; we negate strings by turning the whole string upside down. Double negation elimination now no longer appears as an axiom, theorem, or rule. It can't be stated, or can't be differentiated from  $\underline{p}$  implies  $\underline{p}$ . Of course, we have to be terribly careful. We must not use any symbols that are symmetrical top to bottom, or we won't be able to tell whether they have undergone the negation transformation or not. On the other hand, suppose we use the traditional symbol "v" for conjunction. The effect is that De Morgan's laws need not be stated, are indeed unstatable, being mere fallout from the symbol-typing rules. For example, suppose that you turn "p" and "q" each upside down, put a wedge between them, and then turn the whole string over, thus saying that it is neither the case that not-p nor that not-q. The result, "pvq", is a string that is more naturally read straight off as saying, simply, that p and q. The possibility of swapping symbol-typing rules for axioms or rules of inference is not then an artifact resulting merely from the peculiarities of identity.

Next, I would like to argue that when we turn from a representational system written on paper in a public language to a representational system in the mind or brain, both the distinction between axioms and inference rules and the distinction between typing rules and inference rules tend to break down.

Consider first the apparently nearly self evident truth that although it is possible to build a logical system that has no axioms but only rules, it is not possible to build a system with no rules but only axioms. The rules of a system cannot all be represented explicitly (Carroll 1894). In a traditional formal system, each axiom or hypothesis is written down on a separate line of paper. The system unfolds as new sentences or formulae are derived from these by rule and written down below. In such a system, adding axioms Cwriting down more sentences at the top of the pageC won't by itself determine how these axioms will be used in order to guide derivations within the system.

Rules telling how the axioms, as well as the other strings, are to be manipulated must be given in a metalanguage which the system builder reads and, if appropriately inspired by them, then performs the appropriate transformations.

But if the system is unfolding not on paper but in the head, the "user" is just another part of the head. Since representations in the head are just head-structures designed to vary according to how the world varies, this user part may itself constitute a representation. Taking a childish example, imagine an inference machine designed to perform inferences using universal categorical sentences as major premises and constructed in the following manner. Premises representing that All A's are B's are entered by constructing a sliding board between two ports, the top port being an A-shaped hole, the bottom a B-shaped hole. Sentences ascribing predicates to individuals consist of pieces of putty the colors of which name individuals and the shapes of which

("A", "B", "C") ascribe properties to them. These pieces of putty are gently pushed across the tops of the constructed slides, where they enter the ports through which they fit, proceeding to the bottom where they are pressed down through the lower ports and change their shapes accordingly. They thus become conclusions. All A's are B's and a is an A thus yields a is a B.

Surely this image could be improved on, but the principle should be clear. There is no reason why one premise has to lie passively beside the next in a representational system in the mind or head. There is no reason why the structure, the mechanism, that operates upon a representation during inference may not itself be a representation that has been molded or tuned to perform its appointed tasks, reflecting something in the dynamics, uniformities, or logic of some aspect of the environment, for which aspect it stands. Ways that various individuals' inferencing systems are put together can themselves be representations, so long as they are determined by learning under the influence of individuals' environments, such that variations in ways of being put together correspond, systematically, to variations in environments, according to the design of the learning systems.

If the difference between rules of inference and premises of inference in a cognitive system is not clearly marked, the difference between typing rules and identity axioms or rules for a cognitive system is altogether chimerical. To see exactly why this is so, two more failures of parallel between formal systems and the way representations are used by cognitive systems need to be recognized. One concerns the conventional and public nature of symbol typing in formal systems. The other concerns the use of duplicates as the sameness markers in formal systems.

Earlier I remarked that, typically, nothing is explicitly said about the symbol typing rules when a formal system is laid down, but that it is assumed that these unspoken rules concern parameters or limits on variation in physical form. Better, we pretend that these unspoken rules concern sameness, that is, concern duplication, of physical form. The rules we actually employ are derived, at the start, from irregular rather disjunctive conventions for determining what is the same symbol type. There are, for example, numerous conventional styles of writing the same letter by hand, and numerous type fonts, across which the shape of a particular letter may vary in a rather irregular way. Compare this, for example, with the typing of words, where practices can be quite disorderly. The rules may be quite disjunctive and may include many exceptions. For example, in English, contrasting pronunciations of "schedule" (s-k-e-dule vs. sh-e-dule) count as tokens of the same word type while exactly the same contrast between the pronunciations of "mask" and "mash" or "skin" and "shin" produces different word types. Moreover, the methods we use in practice to determine what counts as another token of the same symbol type concern not just shape but how the person who wrote the symbol intended it to be taken. Recall that most actual formal systems are originally developed in somebody's handwriting on paper and that a typical way of passing them on is by writing on a blackboard. What counts as an "a" or as an "~" then is what was intended to be an "a" or a "~," having been purposefully copied, carefully or carelessly, competently or incompetently, with or without consciously added style, on the model of

earlier "a"s or "~"s.<sup>2</sup> Why do we keep up this pretense that formal symbols are typed by non-disjunctive exceptionless rules on shape?

There is one innocent reason and, I believe, one lingering guilty reason why we do this. The innocent reason is that proofs of consistency and completeness in formal systems work, exactly, by treating systems as if their symbols and WFFs corresponded to well-defined simple physical shapes and well-defined configurations of these. This assumption makes the proofs easy, indeed, possible, and nothing is distorted thereby. The guilty reason is that it is implicitly assumed that the only true marker of sameness in content, the only way sameness of content can be directly represented, is by duplicating representations. That is, the passive picture theory of the act of identifying hovers over. Or it may be assumed that identifying is reacting the same way to what is identified, so that the same physical form will be needed to produce the same reaction again.

But ease in proving consistency and completeness is clearly irrelevant to how cognitive systems work. And the passive picture theory, and the repetition theory of identifying, I have argued, are mistaken. If we keep clear on these issues it becomes evident that there is nothing different in principle going on in a mind that uses "sameness of form" to mark identity if that mind happens to define equivalence classes for these forms disjunctively, or with numerous exceptions, dividing up its space of forms in as gerrymandered a way as you please. All that is needed is that the typing rules used remain consistent with one another. Should the brain mark sameness by equivalence classes of physical type, there is no distinction in principle between systems that mark neatly by perfect duplication of some aspect of form, and systems that mark messily with lists and disjunctions and exceptions. There is nothing magical about simple non-disjunctive typing rules where classes of physical forms mark samenesses.

Now in a public symbol system, which similarities determine that two symbol tokens are of the same type depends on the conventional practices of the language community. That "defence" is read as the same word as "*defense*" and also as the word "DEFENSE," for example, is a matter of public convention. That "Cicero" is not read as the same word as "Tully" is equally a matter of public convention. Certain physical forms and not others are grouped into the same representational type because someone, in this case the general public, reads them that way. And if an individual user of the public language comes along who happens to, or learns to, read "Cicero" and "Tully" as

---

<sup>2</sup> For fuller discussion of this etiological principle in grouping words and other symbols into types, see (Millikan 1984 chapter 4, in press c; Kaplan 1990).

equivalent in typeCthese produce thoughts of exactly the same typeCthat doesn't change the conventional typing rules for words used by the community. It doesn't change the typing rules for the public language.

But for mental representations, there is no distinction like that between public convention and private response. Whatever the individual mind/brain treats as the same mental word again IS the same word again. For mind-language there are no conventions Cthere is only the private user. Nor is there any reason why mental typing should not evolve in an individual mind or brain over time. If the private user changes her habits, then the typing rules for her mental representations will change. This is because the typing rules ARE nothing but her dispositions to coidentify.

For the mind, there also is no distinction like that between an identity axiom or postulate,  $A=B$ , written at the top of the page, and a typing rule. For there is no distinction like that between what is written on the paper and what is written in the structure of the readerCin the structure responsible for conforming the reader's reactions to a certain typing rule. One structure responsible for brain coidentifying patterns is on a par with any other; all are equally "written" in the brain. Write an identity sentence, that is, a structure responsible for producing certain coidentifications, in neuronal patterns instead of in graphite, and the distinction between identity sentence and interpreting mechanism vanishes. Whether the mechanism in the mind effects only that the mental Ciceros's get coidentified with the mental Cicero's, or also that the mental Cicero's get coidentified with the mental Tully's, this mechanism is no more nor less of an extra postulate one way than the other. What is the alternative? That to be an identity postulate it would have, literally, to be physically shaped like this: "Tully=Cicero"? Marking sameness, however that's done, and fixing identity beliefs is exactly the same thing.

Thus if we think carefully about the effects of an equals marker on the system that understands it, the distinction between it and a Strawson marker collapses. What effect are we to imagine mental Cicero  $\equiv$  Tully to have if not, precisely, that it changes the mind's dispositions to mental typing? Henceforth, mental Cicero and mental Tully will behave as representational equals. They become the same mental word, that is, they are ready to be coidentified. But if this is so, the mental equals marker behaves exactly like a Strawson marker. It merges two thought types into one, threatening equivocation in thought, and doling out to each thinker just one mode of presentation per object.

We must conclude, I think, that the peculiar effect of the Strawson markers was on us, on our understanding, not on the operation of the cognitive systems modeled. Systems that use Strawson markers grasp identities by explicitly changing their mental vocabularies, replacing two representations with one. Systems that use equals markers do exactly the same thing but implicitly, changing merely the typing rules for their mental vocabularies, that is, merely the functions of the symbol forms.

#### ' 12.4 Negative Identity Judgments

One possibility concerning negative identity judgments and the undoing of identity markings hasn't been dealt with yet. Imagine a system that keeps a log of the various changes made in the representational system as identities are marked, and keeps a log

of mediate inferences that pivot on these marked representations. Compare the way modern word processing programs can keep track of the last two or three hundred commands carried out. Then if a mistake is discovered, the "undo" button can be pressed until the system is returned to the point where the original false coidentification was made. Different pieces of information that were attached to the same Strawson dot at different times then have different statuses in case of emergency. Indeed, might we say that they represent predicates attached to the same subject but under different modes of presentation? How many of the various purposes of Frege's modes of presentation could differences of this sort serve?

I have not explored these questions because I think such a model is completely lacking in psychological plausibility. Imagine keeping such a log on all the times you have ever reidentified or made inferences about your husband or mother! Of course it is true that were I seriously to suppose, say, that Mark Twain was not Samuel Clemens, I might have some idea how to guess which of my beliefs about this double person should be attached to which name. Certain facts would cohere with Twain's role as an author, others perhaps with his role as public speaker or builder of the Twain house in Hartford. But this untangling would certainly not be done on the basis of a memory of when and in what order I had discovered or inferred what about Twain. It would be done using a theory about how I had got two men so mixed in my mind, and by speculating about which items of information are most likely to have come from which source.

#### ' 12.5 The First Fregean Assumption

How then does the Fregean avoid the Strawson image with its threat of equivocation in thought and its frugal offer of just one mode of presentation per object? The trick is to imagine that how a thought functions has no effect on its content. One assumes that how the mind understands its thoughts is irrelevant to their significance. Throughout I have assumed, on the contrary, that use does affect representational value. I have assumed that what marks content sameness in thought is whatever the cognitive systems read as marking sameness, or what they are designed to read as marking sameness. I assumed that if thought tokens are marked to function as representing the same, this will affect their representational value. In particular, if this marking conflicts with other factors relevant to representational value, say, with the information content of the tokens so marked, or with other ways their associated referents may be determined, then there will be equivocation in content. Sameness is represented yet different things are represented. Visagings of conceptual contents need not be consistent, nor is inconsistency in conceptual content discovered by a priori inspection.

The Fregean view assumes, on the contrary, that insertion of a sameness marker (an identity judgment), hence change in the employment of the marked terms, has no bearing on content. Placing a mental equals sign between mental Cicero and mental Tully has no effect on the representational value of either, even if Cicero is not in fact Tully. Similarly, if duplicated thoughts are in fact thoughts of the same, each token of Cicero referring again always to the same (rather than acting, say, like the English word "he"), this depends in no way on the fact that duplication is what is read by the mind as marking identity. Thought typing is determined independently of thought use.

This Fregean assumption implies, I believe, that thoughts are not mental representations. For we cannot suppose that a representation could be a mental representation, a representation for mind, yet that its representational value was independent of its effect upon mind, independent of how the mind reads it. And, of course, Frege himself did not hold that thoughts are mental representations. Fregean senses are abstract entities that bear their contents quite independently of whether or how a mind "grasps" them. The conclusion that classical Fregean modes of presentation are not compatible with a representational theory of mind is not then a criticism of Frege.

But if we propose to defend any sort of representational theory of mind, we cannot also keep Fregean modes of presentation.

### ' 12.6 The Second Fregean Assumption

According to Frege there are informative and also uninformative identity claims. Uninformative identities are so called because they do not inform us of anything not already immediately known a priori. Presumably these claims also cannot be false. Frege is not supposing that there might be false identities that we cannot help but affirm. The Fregean senses that figure in uninformative identities function psychologically as would thoughts marked with duplicate markers. But on Frege's view this way of functioning is not, of course, what determines that duplicate graspings of duplicate Fregean senses always have the same content. Function has no effect on content. On a representationalist view, sameness markings do force both marked thoughts to refer to the same thing, hence if the markings are wrong, forces both to refer equivocally, but on a Fregean view, the referents of duplicate thoughts are determined independently of the mind's way of being governed by them. That the thinker identifies the referents as one and the same is in no way responsible for them being the same. What is the guarantee, then, that the referents of duplicate thoughts actually ARE the same? (Or if what you mean by "duplicate thoughts" includes that they have the same referent, what is to guarantee that the mind that grasps two thoughts can tell whether these thoughts are indeed duplicates?) How can there be uninformative identities that are at the same time certain to be real identities and not merely false appearances of identity?

This line of questioning highlights the internalist assumption built into the Fregean position. What is duplicated when "the very same thought" is repeated must be something that is simultaneously (1) compelled always to bring with it the same referent and (2) capable of being unmistakably known by the mind, when the mind duplicates its entertainment, it as being the very same thought. That, I take it, is one role of a Fregean sense: it always determines the same referent regardless of the context, the grasper (understander), or the use, and its identity is transparent to mind.<sup>3</sup> More generally, that

---

3 If it were not transparent to mind, then that one cannot think a contradiction about a thing while thinking of it under just one mode of presentation could not be criterial of sameness of sense, nor could it be assumed that uninformative identities are never false identities. (If Frege's position is not that identity of sense is transparent to mind, certainly this is what many have thought his position to be. The purpose here is not, of course, Frege exegesis but clarification of where certain incompatibilities of position lie.)

which completely determines the referent must be exactly the selfsame as that which, when duplicated, constitutes a grasp of sameness. Otherwise the appearance of sameness might not be veridical.

It follows that whatever determines the reference must be entirely internal to mind.

Reference cannot be affected at all by, say, the external causes of thoughts, or their natural informational content, for there can be no certain internal or *a priori* mark proving the external causes or the informational contents of two thoughts are really the same.

' 12.7 Rejecting Identity Judgments.

I have been arguing in this book for an externalist and representationalist position on thought, but it is crucial that we not rely on a language-of-thought model of mental representation. This misleading model offers an image of thought as having two levels of sameness marking, on one of which sameness of content for concepts is not an empirical issue but a matter of mere inner form. It pictures errors about identity as impossible on this level and, correlatively, pictures the difference between valid and invalid mediate inference as though it were distinctly marked from within. But on externalist grounds, no distinction can be clearly marked in thought between valid inference that relies on false identity premises and invalid inference. There is no analytic/synthetic distinction for identity as grasped in thought.

It follows, for example, that the Quinean distinction between tautologies and merely analytic sentences and, similarly, the distinction customarily drawn between inferences valid due to logical form and those whose validity depends on the meanings of nonlogical terms, cannot be drawn for thought. In a similar vein, we have a linguistic convention that any adjective grammatically marked as comparative in English, for example, by adding the suffix "-er" expresses a relation that is transitive. This gives the appearance that thoughts expressed in the argument form "A is öer than B and B is öer than C, therefore A is öer than C" are valid a priori. But, as Hemple showed us clearly, using the geologist's "harder than" as his example (' 7.2), particular applications of this form are not valid a priori. They are valid only in case the English language happens to conform, in particular cases, to the convention, and conformity cannot be guaranteed, exactly because the transitivity of an empirical relation cannot be known a priori. The linguistic form of the argument makes it valid by linguistic convention, but there can be no guarantee that the convention is manifested in particular cases. Similarly, Whitehead claimed that it is always an empirical matter, in the particular case, that one plus one equals two. One raindrop plus one raindrop sometimes equals one raindrop, and one quart of water plus one quart of alcohol equals less than two quarts. The convention is to use numbers and numerical operators only where number theory applies, but that it applies in particular cases is known empirically.

Putting the inevitably aposteriori nature of our grasp of sameness in John Campbell's terms (1987/88), there is no such thing as completely "manifest sameness of reference". The mental sentence image causes us to overlook the most central fact about cognition, namely, that its most difficult job is to get the empirical identities right, to create a coherent, nonredundant and nonequivocal mental representational system.

Without such a system, or something sufficiently close, there can be no conceptual thinking at all.

There can be no representation of sameness in thought without sameness marking, and there are no substance concepts without representings of sameness. Indeed, what substance concepts are initially for is grasping, which requires somehow marking, sameness in substances. But since there also is no difference between marking sameness and fixing identity beliefs, it follows that there are no representations of substances that are free from the possibility of empirical error. It is always contingent that a substance concept represents univocally, or represent at all. We must proceed very carefully here, however. It does not follow that substance concepts somehow make claims, or that they are "theories," or that they really are identity judgments in disguise.

There are lots of ways to do things right rather than wrong without making claims or holding theories. You don't make claims when you stand up to walk just because it's possible you could trip and fall. Similarly, you don't make claims when you develop substance concepts or when you mark identities in thought. Erroneous identification is not failure on the level of know-that but failure on the level of know-how. It is failure in an activity. Standing back from a failed activity it is often possible to explain its failure by pointing to some proposition that, had it only been true, would have prevented the failure. Had that wrinkle in the rug not been there, I would not have tripped. It doesn't follow that my attempt to walk involved a judgment that no wrinkles were in the rug. Similarly, when representations carrying information about different substances are wrongly coidentified, it is true that had they carried information about the same they would not have been coidentified wrongly. It does not follow that the act of coidentifying, or of identity marking, is a judgment that they carry information about the same. It is not a judgment about sameness of content.

Rather than substance concepts being implicit judgments or theories, it is better to say that, as distinguished from an identity sentence or assertion, there is no such thing as an identity judgment. It is not the job of an identity sentence to induce a belief. Its job is to induce an act of coidentifying. True, an identity sentence has a grammar superficially like that of ordinary subject-predicate fact stating sentences. It has traditionally been recognized, however, that it does not have a logical subject and a logical predicate. Only if we insist on modeling thought on language should there be a temptation to assimilate what an identity sentence produces Can act of coidentifyingC to what a subject-predicate sentence characteristically produces Cnamely, an intentional attitude.

Grasping an identity is not remotely like harboring an intentional attitude. Similarly, mistaking an identity is not harboring a false belief. It is an error of its own kind. Misidentifying is not, in central cases, an innocent act of false judgment, but an act that tends to muddy the very content of the thought involved, corrupting the inner representational system. The development and maintenance of relatively clear and distinct ideas is a substantive ongoing activity. Descartes was quite right that not all ideas are clear and distinct, indeed, that some are materially false. He went astray only in failing to see how much more than mere armchair reflection is involved in the activity of

clarifying our ideas (Chapter 7).

### ' 12.8 Rejecting Modes of Presentation

Where Frege drew one distinction, between sense or mode of presentation and reference, there seem really to be two distinctions, or anyway two phenomena, confused together. First, there is the distinction between concept and conception, that is, between designating conceptual abilities by their ends only and designating them by their means (or by certain of their means) as well. Second, there is the possibility of having more than one concept of the same, these concepts being separate and not marked as of the same.

This second possibility gives rise to an embarrassment in terminology. I have spoken of two different ways to understand the notion "same ability" hence "same concept," depending on whether "the same" means the same end achieved or whether it means the same end achieved by the same means (' 1.9, ' 6.3). But it appears now that there is also a third way to interpret the notion "same concept." If concepts can be called "the same" when they involve abilities to identify the same, thus allowing you and me to have many of "the same" concepts, then it will also be possible for me to have two different tokens of the "same concept" because I have failed to co-identify two concepts of the same thing. "The same" concept token, is identified neither by its end alone nor by its end plus its means. If a person harbors two different tokens of the same concept these will, of course, be governed by different conceptions, by different means, but that is not what makes them two. Similarly, someone might be said to have two separate abilities to achieve the same end, where the point is not merely that they know several ways to achieve that end. Suppose, for example, that before it was known that scarlet fever is the same disease as rheumatic fever but in a different form, there was a doctor who knew how to cure scarlet fever by one means and knew how to cure rheumatic fever by another. Such a doctor might be said to have two different abilities to cure strep infections, two tokens of the same ability.

In sum, we should distinguish, first, among concept tokens, and second, among two different sorts of concept types, types distinguished by ends and types distinguished by ends plus means. It is also possible to classify concepts more abstractly, according merely to some of their means. But here we should proceed very cautiously

One of the things we have put in place of Fregean modes of presentation is conceptions. I have described conceptions as the "means" by which the thinker knows how to reidentify a substance. Is the word "means" in this usage a singular noun, a plural noun or a mass noun? Can we, for example, sensibly speak of "some of the means" of a concept? In ' 11.5 I argued that in the case of recognitional abilities, the means actually or possibly used to reidentify a substance cannot be divided into discrete countable "ways of identifying." "Means," in this context, seems to be a mass noun. Conceptions don't for the most part divide into discrete parts. The individual recognitional abilities that any two people have have, allowing them to reidentify the same substances, are very unlikely to be identical in means, nor are they likely to be describable in a readily available way.

On the other hand, some features of these means may be describable. For

example, I might be able to recognize Xavier by sight but not by his voice on the phone, or recognize lemons by smell but not by their name in Russian, or recognize Jon Jones by the fact that he is the one chairing the meeting but not by scanning his face. Of course, that I recognize Xavier by sight does not tell how I recognize him by sight, nor does my recognizing Jon Jones by his chairing the meeting tell how I recognize which person is chairing. Mentioning certain features of the conceptions I use is not telling the whole story about how I recognize, even the whole story on some specific occasion. Still, it is possible, sometimes, to describe conceptions by aspects of their means in this rough sort of way, and this allows us to give certain kinds of psychological explanations.

Where a person's reidentifying abilities have proved fallible, descriptions of intentional attitudes referring only to the objects and properties these concern, that is, purely referential descriptions of them, can be misleading. Where it matters to psychological explanation, the normal assumption is that relevant identities have been correctly recognized. If they haven't, or easily might not have been, we describe thoughts by reference to rough aspects of the conceptions involved. Thus we explain why Paul didn't speak to the woman he admires, even though she was present and he very much wanted to, by the fact that he didn't know what she looks like. Similarly, we may refer to identifying knowledge that forms part of a person's conception of an object, or make reference to a name by which they recognize an object. In this way we may move back and forth between purely "transparent" descriptions of a person's intentional attitudes and somewhat more opaque ones. Consider, for example, a case with which I opened this book. It would certainly be misleading to say, without further explanation, that someone at the the Yale Alumni Association headquarters wished to know whether I knew where I was, but not misleading to say that they wished to know whether Mrs. Donald P. Shankweiler knew where Ruth Garrett Millikan was. From the fact that the conceptions governing the two different concepts of me could be described individually in this manner it does not follow, however, that the entire conception governing either concept could be described. Nor does it follow that there could be a description of an intentional attitude that was entirely opaque, making no transparent reference to any objects or properties outside of the thinker.<sup>4</sup> Even knowing someone merely by their name is not having a conception of them describable in completely opaque terms. "The ability to recognize the name X" is a transparent description of that ability. So-called "opaque" descriptions of conceptions and intentional attitudes are never more than semi-opaque.

Besides semi-opaque descriptions of various means supporting real substance concepts, there also can be, of course, semi-opaque descriptions of various "would-be" means but that fail to support real substance concepts. That is, unbeknownst to their possessors, they are not the means for any real recognition abilities. These substance concepts are "empty" or, more accurately, they are not substance concepts at all. An

---

4 Various different ways of describing intentional attitudes are discussed in (Millikan 1984) Chapter 13.

ability that is not an ability to do anything is not an ability at all. Empty substance concepts result from failures of the mechanisms designed to develop substance concepts. They are "concepts" only in that their biological purpose was to have been concepts. Nor should they be confused with concepts we merely pretend to have, such as the concept we pretend to have of Santa Clause after we are grown. Semi-opaque descriptions can also be given of pretend concepts, of course, but that is another matter.

Opaque descriptions characterize aspects of conceptions, that is, aspects of ways of identifying substances. They do not describe ways of thinking of substances. "Via a definite description" and "via a proper name" are not ways of thinking of things. Nor, of course, is "by recognizing her face" or "by recognizing her voice" a way of thinking of a person. A neo-Fregean tradition has it that perceiving a person is a sort of way of thinking of them, indeed, it is supposed, an indexical way of thinking of them—thinking of them via an indexical mode of presentation. This, I believe, is a serious confusion. Suppose that you see Alice and track her perceptually, picking up information about her as you proceed but without recognizing her as Alice. Then for the moment you have two concept tokens of Alice that you have not co-identified. Your current tracking ability supports a "naive Strawson mode of presentation," if we may still use that terminology, which is separate from the naive Strawson mode in which your prior knowledge of Alice is stored. As we saw in § 11.2, however, this mode is no Fregean mode of presentation. Furthermore, once you have recognized Alice, no distinction of "modes" of any kind remains.

To be sure, your current perception of Alice yields, in part, a special kind of information about Alice, namely, information about her current spatial relation to you. But that it yields information about her relation to you has nothing to do with indexicality. This is a subject I have addressed at length elsewhere, and will not pursue further here. There are, I have claimed, no mental indexicals at all, least of all any (so-called) essential indexicals (Millikan in press a).