

Fachgruppe Sprachwissenschaft

Universität Konstanz



Arbeitspapier 79-6

Proceedings of the Konstanz Workshop
„Reference and Anaphorical Relations“
Klaus von Heusinger & Urs Egli (eds.)
November 1996

Reference and Inference

Jaroslav Peregrin

Reference and Inference

Jaroslav Peregrin^{*}

Summary

This paper discusses the relationship between the concept of reference and that of inference; the point is to indicate that contrary to the usual view it may be good to see the former as ‘parasiting’ on the latter, not the other way around. The paper is divided into two parts.

In part one, I give an (unsystematic) overview of the development of logical tools which have been employed in the course of the analysis of referring expressions, i.e. definite and (specific) indefinite singular terms, of natural language. I present Russell’s celebrated theory of definite descriptions which I see as an attempt to explain definite reference in terms of unique existence (and reference in general in terms of existence *simpliciter*); and I present Hilbert’s ϵ -calculus as an attempt to explain existence in terms of choice. Then I turn to contemporary, dynamic approaches to the analysis of singular terms and point out that only within a dynamic framework can the Russellian and Hilbertian ideas yield a truly satisfactory analysis of singular terms, and consequently of reference and coreference. I call attention to the fact that current results of formal semantics demonstrate the advantages of viewing singular terms as denoting *updates*, i.e. as means of changing the context (information state), and especially that part of the context which I call the *individuary*.

In part two I then turn to the discussion of the nature of such explications; especially to the question whether it forces the acceptance of a representational view of language. I answer the question negatively; I deny that we should see discourse representations, information states, or individuaries, which play the central roles within contemporary semantic theories, as descriptions of a mental reality; I try to show that these entities can, and indeed should, be seen as tools internal to our accounts for singular terms’ inferential capacities. Therefore I conclude that we should not take the obscure concept of reference at face value, but rather as parasitic upon the clear concept of inference.

^{*} I am grateful to Klaus von Heusinger for valuable comments and criticism.

1. The Logical Grip on Reference

1.1 Russell

Some expressions of our language are seen as doing their linguistic jobs by referring to definite things of our world. How do they manage to do this?

The classical analysis of *definite descriptions* (and of the English definite article, which is their linguistic hallmark), as presented by Russell (1905), consists in explicating definiteness in terms of unique existence. To say that the king of France is bald is to say that there is one and only one entity which is the king of France, and that this entity is bald. Thus, the sentence

The king is bald (1)

should, according to Russell, be construed as

$\exists x (\mathbf{K}(x) \ \& \ \mathbf{B}(x) \ \& \ \forall y (\mathbf{K}(y) \rightarrow (y = x)))$ (1')

This reconstructs the definite article as a ‘syncategorematic term’, i.e. as something which is not itself a fully-fledged constituent of the sentence. However, if we utilize stronger formal means than those entertained by Russell (which were in fact the means of the first-order predicate calculus), we can reconstruct *the* ‘categorematically’ (in the sense of granting it its own denotation) without violating the spirit of the Russellian analysis.

Thus, helping ourselves to the mechanism of lambda-abstraction, we can rewrite (1') as

$\lambda f \exists x (\mathbf{K}(x) \ \& \ f(x) \ \& \ \forall y (\mathbf{K}(y) \rightarrow (y = x))) (\mathbf{B})$

and this further as

$\lambda g \lambda f \exists x (g(x) \ \& \ f(x) \ \& \ \forall y (g(y) \rightarrow (y = x))) (\mathbf{K}) (\mathbf{B})$

This yields us a formula consisting of three parts which may be put into the natural correspondence with the three components of the analyzed sentence; the definite article thus gets formalized as $\lambda g \lambda f \exists x (g(x) \ \& \ f(x) \ \& \ \forall y (g(y) \rightarrow (y = x)))$, i.e. as a function which takes sets into sets of sets; or, probably more perspicuously, as a relation between sets. The relation holds between two sets iff the first of them is a singleton and has a non-empty intersection with the second one. This is the analysis which has become standard since Montague (1974); and which has given rise to the so called *theory of generalized quantifiers* (viz Barwise & Cooper, 1981).

Alternatively, we can make the following consideration. First, assume that $\exists x (\mathbf{K}(x) \ \& \ \forall y (\mathbf{K}(y) \rightarrow (y = x)))$ is true, i.e. that the extension of \mathbf{K} is a singleton. Under

such assumption, if we denote the single element of the extension of \mathbf{K} as \mathbf{c} , the whole formula becomes equivalent to $\mathbf{B}(\mathbf{c})$. Next, assume that $\exists x (\mathbf{K}(x) \ \& \ \forall y (\mathbf{K}(y) \rightarrow (y = x)))$ is *not* true, i.e. that the extension of \mathbf{K} is either empty or contains more than one element. Then the whole formula is patently false, i.e. equivalent to \perp . This means that if we were able to define a (second-order) function \mathbf{F} which maps singletons on their single elements, and all other sets on something of which \mathbf{B} is inevitably false, we could rewrite (1') as $\mathbf{B}(\mathbf{F}(\mathbf{K}))$; or, writing $\iota x \mathbf{K}(x)$ instead of $\mathbf{F}(\mathbf{K})$, as $\mathbf{B}(\iota x \mathbf{K}(x))$.

The most straightforward way to reach such a function would be to stipulate an object of which everything is necessarily false and to let this object be the value of the function for all non-singletons; or, which is the same, to let the function be defined only for singletons, and to stipulate that any predicate applied to a term which lacks denotation yields a false sentence. In the latter way, we would reach an analysis which would be clearly effectively equivalent to the previous one, which treated the definite article as a generalized quantifier. However, a modification suggests itself: we can also stipulate that a predicate applied to a denotationless term yields a sentence which is not false, but truthvalueless - in this way we can clearly accommodate the idea of an existential presupposition associated with a definite description (as urged by Strawson, 1950).

Russell himself claimed that definite singular terms which are not proper names (*the king of France*) are not in fact referring expressions at all but that they, when properly logically analyzed, give rise to a certain quantificational structure: a sentence consisting of a definite description and a predicate, according to him, says that there is one and only one object satisfying the description and such that it has the property expressed by the predicate (*viz* (1')). Similarly he claimed that a sentence consisting of an indefinite description (*a king*) and a predicate says that there is an object satisfying the description and having the property expressed by the predicate; thus, the adequate analysis of (2) is, according to him, (2').

A king is bald (2)

$\exists x (\mathbf{K}(x) \ \& \ \mathbf{B}(x))$ (2')

In this way, the working of a definite singular term gets reduced to the unique existence of the corresponding referent; and that of an indefinite one to the existence *simpliciter*. We have seen that allowing ourselves of more powerful logical means than Russell himself, we can extrapolate Russell's analysis to explicate definite singular terms via taking the denotation of the definite article to be a function mapping singletons on their unique elements (and nonsingletons on some kind of "nothing")¹. However, no analogous straightforward explication is available for the meaning of the indefinite article, and hence for reference in general.

¹ Russell himself introduced the ι operator as a mere notational short-cut governed by a contextual definition, not as a fully-fledged term (see Russell and Whitehead, 1913).

1.2. Hilbert

Russell's treatment of definite descriptions illustrates the intuitive intimate tie between definiteness and choosability: definiteness, i.e. unique existence, turns out to be a matter of unique choosability. This seems to invite generalization: why not see existence in general as choosability in general? The idea is that something exists if and only if it can be chosen ("picked up"); not necessarily by a particular human subject (whose capacities to really carry out the choice could be limited), but "in principle", or "by God". That there is an F means that an F can be chosen. If we render the possibility of choosing as the existence of the corresponding choice function, we can say that the existence of an item is tantamount to the existence of the appropriate choice function: to say that there is an F is to say that there is a choice function which chooses an F .

Two kinds of objections can be raised against the identification of the existence of an object with the existence of a choice function choosing the object. First, there is the "constructivistic" objection claiming that to be choosable is always more than to merely exist - that claims about the former necessarily violate bivalence, while those of the latter do (*ex definitio*) not. Then there are the scruples of set-theoreticians to turn the intuition directly into cash by embracing the axiom of choice. The first objection can be turned into a purely terminological matter: to say that something exists is to *acknowledge* that it exists and in this sense to *choose* the thing from among other things. Thus, in this sense, if it makes sense to speak about *existence* of a thing, it makes the same sense to speak about that thing being *chosen* - although we have to keep in mind the broad ("bivalent") sense in which the term *choice* is being used. The second objection invokes the well-known set-theoretical perplexities swarming around the axiom of choice, which we are not going to discuss in this paper.²

One of the possible ways to develop this idea is to stipulate the reduction of the axioms of existence to the axiom of choice; this development was carried out by Hilbert. To see the semantic point of the enterprise, let us consider functions mapping nonempty sets on their elements: a function f is called a *choice function on the set U* iff the domain of f is included in the power set of U and $f(s) \in s$ whenever s is not empty. If M is a model, then the set of all total choice functions on the universe of M will be denoted as CHF_M . It follows that if $c \in CHF_M$ and if s is a subset of the universe, then $c(s) \in s$ iff s is nonempty, i.e. iff there is an element of s . If s is the denotation of a unary predicate F ($\{F\}^M = s$), then $c(s) \in s$ iff there is an F , i.e. iff $\exists x Fx$:

$$\|\exists x Fx\|=1 \text{ iff } c(\|F\|) \in \|F\| \text{ for an arbitrary }^3 c \in CHF_M$$

² However, notice that, as Lavine (1994, p. 104) points out, the axiom of choice is problematic only under a certain specific notion of set, namely under the "logical" notion, claiming that items can be collected into a set only if they can be delimited by a criterion.

³ The definition of CHF_M clearly guaranties that this holds for *at least one* $c \in CHF_M$ if and only if it holds for *every* $c \in CHF_M$.

If we now understand ε as denoting an arbitrary function from CHF_M in such a way that εxFx denotes the value of ε for the set, then the following formulas will be clearly valid:

$$\mathbf{F}t \rightarrow \mathbf{F}(\varepsilon xFx)$$

$$\exists x \mathbf{F}x \leftrightarrow \mathbf{F}(\varepsilon x \mathbf{F}x)$$

$$\forall x \mathbf{F}x \leftrightarrow \mathbf{F}(\varepsilon x \neg \mathbf{F}x)$$

Hilbert (1925) showed that if we accept the first of these as an axiom characterizing the ε -operator, we will be able to prove the other two and hence justifiably reduce quantification (and especially existence) to choice⁴.

In this way the ε operator might seem to be capable of explicating existence *simpliciter*, and consequently the semantics of the indefinite article, in a way parallel to that in which the ι operator explicates unique existence and hence the definite article. In one sense, this is indeed the case⁵; however, we should realize an essential difference between the two operators: while ι can be seen as a logical constant denoting a definite function (namely the function which maps singletons on their unique elements and is undefined for non-singletons) and thus furnishing a definite model-theoretical explication of the meaning of the definite article, ε is in fact more like an extralogical constant, ranging over the whole set of choice functions and thus *not* furnishing any definite explication of the meaning of the indefinite article. Besides this, one of the unmistakable features of the articles is their interplay: if *the F* follows *an F*, then the two noun phrases are in the typical case coreferential; and this is something which is not directly reflected by the Russellian and Hilbertian treatment.

Therefore the direct exploitability of the ideas of Russell and Hilbert seems to be limited; and my opinion is that to progress we must ‘go dynamic’ - we cannot have an explication of the meaning of the indefinite article (nor indeed a satisfactory explication of that of the definite one) until we start seeing meanings as “context-change potentials”.

1.3. Dynamic Semantics

To do this, let us first return to the Russellian analysis: it is clear that as it stands, it is *not* adequate to explicate our everyday use of the definite article. Saying *the F* does not usually involve claiming that there is one and only one F, but rather that there is one and only one “salient” F. With respect to this point two principal candidate ways to amend the analysis seem to emerge. The first possibility is to retain the Russellian analysis as such and to retract the

⁴ Hilbert’s treatment of the ε operator was purely axiomatic; he did not consider any kind of model theory. Nevertheless, choice functions clearly represent the straightforward way to put his ideas into the semantic cash. (Notice, however, that Hilbert’s axiom allows for ‘intensional’ choice - the value of εxFx may differ from that of $\varepsilon x\mathbf{F}'x$ even if \mathbf{F} and \mathbf{F}' are coextensional. See, e.g., Meyer Viol, 1995).

⁵ In fact, it seems to be this idea which lays the foundations of Egli’s and von Heusinger’s exploitation of the Hilbertian ideas for semantic analysis (see their contributions in Egli & von Heusinger, 1995).

assumption that all evaluations take place with respect to the general, allembicing universe. This is to assume that the evaluation of a particular sentence may be based on a local, restricted universe, which is the result of the (“pragmatic”) circumstances, in particular of the preceding discourse. Thus, we keep assuming that the sentence *The king is bald* implies that there is one and only one king - not, however, in the general universe, but rather in a local universe determined by the context in which the sentence is being uttered. Hence it implies not that one and only one king exists, but rather that one and only one king is *salient*. Alternatively, the second possibility then is to assume that the Russellian analysis itself must be amended, that the choice function represented by the definite article is defined not only for singletons, but instead that it is capable of using strategies to successfully pick up an element even from some other nonempty sets.

In both cases the analysis rests on such or another formalization of the notion of context. In a context, some objects of the universe are salient, while others are not. The first approximation of the formalization of the concept of context could be thought of as simply a set of objects, a subset of the universe. This kind of context is utilizable by both of the above mentioned strategies: in the first case, we apply the Russellian analysis not to the universe, but rather to the context-delimited subset; in the second case we let the Russellian operator choose not *the only element* of the extension in question, but *the only element which is in the context-delimited subset*.

Taking the concept of context seriously entails subscribing to some version of the dynamic view on semantics, as proposed by a number of semanticists. It leads to reconstructing meaning as resting on (or at least involving) a kind of “context-change potential”⁶, of mapping of contexts on contexts. Contexts may be, and indeed are, captured in a variety of ways - for the problems discussed here it is nevertheless vital that a context somehow “contains” a class of (salient) individuals. In this way we can reconstruct definiteness not as presupposing unique existence, but rather as presupposing unique “referential availability”.

To sensitize the denotation of a sentence to context, the denotation must become a “context-consumer”, it must become functionally dependent on the context (formalized in such or another way). For the semantics to become really dynamic, we must turn denotations into not only “context-consumers”, but also “context-producers” - so that an utterance might consume a context produced by a preceding one. From the point of view of anaphora it is again vital that the context somehow contains individuals - so that one utterance can introduce (raise to salience) an individual, to which other utterances may then refer back. Our current interests lie only in that part of the context which functions as such an individuals-container (which other parts a context can or should have is not a theme for us now) - to have a name for it, let us introduce the term *individuary* (devised by analogy to the well-established term *bestiary*).

⁶ As far as I know, this term is due to Irene Heim. See. e.g., Heim (1992).

In the simplest case, an utterance introduces a single object by using an indefinite description, and a subsequent utterance uses a definite description, or a pronoun, to pick it up. This is the case of

A man walks. He whistles. (3)

For this example, it suffices to construe the individuarity as a single slot which can be occupied by an object, or be empty. The phrase *a man* fills the slot with a (“fixed but arbitrary”) individual, and the phrase *he* is then interpreted as referring to this very individual. (There are essentially two ways of formalizing this “fixed but arbitrary”: we can either represent the assignment of such an “arbitrary” object by means of a whole class of assignments of “real” objects, or we can introduce some kind of genuine “arbitrary objects”, pegs, which are capable of being identified with real objects. The former is the way of Groenendijk and Stokhof, 1991, the latter is that of Kamp, 1981, and also of Groenendijk, Stokhof and Veltman, 1996).

The situation is more complicated once there is more than one object, which is so passed from one utterance to another. Take

A man meets a woman. He greets her. (4)

In this case, when it comes to the anaphoric reference, it is necessary to choose the right referent from among more salient objects. This is not to say that there is *always* the right choice, but in many cases, such as in this one, there clearly is. This possibility of the right choice implies that the items in the individuarity, the salient individuals, have to be in some way characterized, for it is only on the basis of some characteristic specification that we can distinguish between them and carry out the choice.

1.4. The Structure of the Individuarity

Usual approaches to dynamic semantics take the resolution of anaphora to be a matter of “coindexing”; they in fact assume that the ‘real’ semantic analysis begins only *after* coreference has been settled. Thus, they assume that getting the right semantic analysis of (4) is the matter of being given (4a), and not, say, (4b), as the input of the semantic analysis.

A man₁ meets a woman₂. He₁ greets her₂. (4a)

A man₁ meets a woman₂. He₂ greets her₁. (4b)

Under such an approach, the salient item stored into the individuarity is characterized in such a way that it is associated with an index, or with some other formal item working to the same effect (Groenendijk & Stokhof’s *discourse marker*) - the individuarity can thus be constructed as a unique assignment of objects to indices.

However, in this way the semantician gets rid of a part of the burden which is evidently *his own*. It seems obvious that the semantic analysis resulting from (4a) is the *right* one, while that resulting from (4b) is a *wrong* one - and this comes from understanding (4) - hence, it is the matter of the *semantics* of (4) and it has to be brought out by the semantic analysis of (4). Thus, the semantician should not wait for someone to pass him a coindexing, he should aim at yielding the right analysis directly. And this requires a more *substantial* characterization of the items in the individuary.

The idea that comes to mind is to store the individual with the “attribute” which is employed to introduce it: to store the item that is raised to salience by means of the phrase *a man* with the attribute *man*, and, more generally, that raised to salience by means of the phrase *a(n) N* with the attribute *N*. This would turn the individuary into an assembly of the <individual,attribute> pairs - a phrase *the N* would then look for the individual which is paired with the attribute *N*. Thus, the first sentence of (4) would fill the individuary with the pairs <I₁,*man*> and <I₂,*woman*>; and for the resolution to succeed it would be enough to secure that the pronoun *he* seeks an individual with the attribute *man* (i.e. it is in effect equivalent to *the man*) and the pronoun *she* seeks one with the attribute *woman* (it is equivalent to *the woman*)⁷.

The question now is how to construe the word ‘attribute’. We may take attributes simply as *words* - and thus form the individuary as a collection of individual-term pairs. However, this would block the resolution in intuitively clear cases, where one uses a slightly different word, like

A man walks. The guy whistles.

This may lead us to abandon taking attributes as terms in favor of taking them more as *meanings* of terms - where meanings can be, in turn, taken to be extensions, intensions, or something else. (Then, however, we may have to face the opposite problem: the problem of “overgeneration” of anaphora resolutions. In general, the optimality of the account for anaphora is the matter of fine-tuning the fine-grainedness of the attributes - and due to the heterogeneity of language there is little hope that we can find one universally optimal solution⁸.)

Anyway, within the dynamic approach, both the meaning of an indefinite noun phrase and that of a definite one get explicated as *updates*, as means of inovating the current context and especially its individuary. An indefinite noun phrase changes the individuary by introducing a new inhabitant characterized in a certain way; a definite noun phrase does not change the individuary (in this sense it is a *trivial* update, a ‘test’), but searches it for the existence of an individual with a certain specification, thereby triggering a presupposition that such an individual is indeed present there, i.e. that it is “referentially available”. This opens the possibility of adequately explicating the meanings of the indefinite and definite articles: both get

⁷ The most straightforward way of exploiting this idea is perhaps Heim’s (1982) *file change semantics*.

⁸ This is analogous to the case of objects of propositional attitudes: they also seem to be sometimes like propositions, whereas sometimes rather like sentences.

explicated as functions which map properties on corresponding updates. Thus, the denotation of *a* is the function which maps the denotation $\|N\|$ of the common noun *N* on the update which stores $\langle I, \|N\| \rangle$ into the current individuary and makes the whole singular term refer to the *I* thus introduced; and the denotation of *the* is the function which maps $\|N\|$ on the update which searches the current individuary for an $\langle I, \|N\| \rangle$ and, if successful, makes the whole singular term refer to the *I* thus found.

One of the possibilities is to take the attributes as sets of potential referents (this is straightforward if we stay on the level of extensions; but attributes can be constructed as sets of individuals even when we embrace intensions - in this case they are sets of not only actual, but rather also possible, individuals⁹). In that case, individuary is explicated as a choice function taking sets of individuals into their members - this is the framework introduced by Peregrin & von Heusinger (1995). In this setting, *choosing the right referent* is reconstructed as bringing the choice function to bear on the relevant subpart of the class of potential referents.

⁹ It is not without interest to note that this is in fact the way in which the concept of intension was approached by Rudolf Carnap (see esp. Carnap, 1955), who is responsible for its current dissemination.

2. Reference as Resting on Inference

2.1 The Nature of the Individuary

I consider one of the most important questions concerning this whole enterprise to be: what are the individuarities, and more generally the ‘discourse representation structures’ of all sorts, supposed to be? Few people practicing dynamic semantics seem bothered by the question: they apparently assume it to be straightforwardly answerable in terms of ‘mental representations’ or ‘cognitive states’. I do not think they are right: I am convinced that to explain the linguistic by means of the mental is to explain the clearer by the more obscure; and, moreover, it is to block the requisite possibility of going the other way around, namely to use the linguistic to account for the mental. Therefore, I want to propose an alternative answer: the answer that the individuary is our (i.e.: our theoreticians’) way of accounting for the inferential properties of anaphoric expressions.

This question is also essentially relevant for the proper understanding of the analysis of definiteness in terms of choice outlined above: if we see the apparatus of choice functions as descriptive of mind or cognition, we will be likely to see choice functions as reports of actual ‘mental actions’ carried out by speakers and hearers (and we will be likely to pose questions such as *how* they carry out the relevant choices or *why* they choose as they do and not otherwise); whereas if we see it as our way of accounting for certain valid inferences, no such questions really make sense. In the former case, ultimate answers will appear to be buried within people’s heads and we will have to set upon the slippery path of introspection; whereas in the latter case we shall be able to rest on the relatively solid notion of inference as based in the norms of verbal behavior.

So, how can we characterize the inferential behavior of an expression? Of course by stating some characteristic inferences in which it features; and then perhaps by assigning it some kind of ‘value’ which in a certain sense ‘encapsulates’ its inferential behavior as fixed by them. The paradigmatic example is capturing the inferential role of a propositional connective like **&**: this, as is well known, can be captured by stating that we can always infer $X\&Y$ from X and Y , and that we can always infer X , as well as Y , from $X\&Y$; i.e. by the following ‘introduction’ and ‘elimination’ rules¹⁰:

¹⁰ For the general notion of introduction and elimination rules see, e.g., Prawitz (1965).

$$\begin{array}{l} X \quad Y \\ \hline X \& Y \end{array} \quad (\&I)$$

$$\begin{array}{l} X \& Y \\ \hline X \end{array} \quad (\&E_1)$$

$$\begin{array}{l} X \& Y \\ \hline Y \end{array} \quad (\&E_2)$$

The same can be clearly achieved by letting & denote the usual binary truth function mapping the pair $\langle truth, truth \rangle$ on *truth* and any other combination of truth values on *falsity*.

Now what are the ‘inferential properties’ of singular terms? Could we formulate something like the introduction and elimination rules for them? First, let us look at pronouns; because they are, from our viewpoint, particularly simple. The basic inferential characteristics of a pronoun like *he* is the inferability of *X Ps and Qs* (*John walks and whistles*) from *X Ps and he Qs* (*John walks and he whistles*) and vice versa. To articulate this intuition formally, let us take a very simple fragment of the predicate calculus, a fragment containing only individual constants, unary predicate constants and the operator &. Let us consider the addition of a new constant, say \leftarrow , functioning like a ‘universal’ pronoun; this means that $P(a) \& P'(\leftarrow)$ is always equivalent to $P(a) \& P'(a)$, i.e. that we can characterize \leftarrow by the following natural deduction rules¹¹

$$\begin{array}{l} P(a) \& P'(a) \\ \hline P(a) \& P'(\leftarrow) \end{array} \quad (\leftarrow I)$$

$$\begin{array}{l} P(a) \& P'(\leftarrow) \\ \hline P(a) \& P'(a) \end{array} \quad (\leftarrow E)$$

¹¹ From the purely logical viewpoint, these rules, and consequently \leftarrow , are clearly not of great interest, for what they introduce is in fact nothing but certain notational variants of certain conjunctive statements.

It is clear that \leftarrow cannot be taken as simply an individual constant: its inferential role is more complicated than as to be model-theoretically represented by the assignment of a single referent. Informally, we can describe the inferential role by saying that \leftarrow “behaves like a if it follows a clause containing a , it behaves like b if it follows a clause containing b etc.”; thus we may specify its referent only relatively to the referent of the preceding individual constant. Hence we may come to think of assign \leftarrow something like the identity function from individuals to individuals.

This idea can be put to work by making the semantic values of statements into functions from individuals to individuals, and letting $\&$ work as a concatenation in the following way¹². First, we change the semantics of our fragment of the predicate calculus (without \leftarrow) by letting statements denote mappings of a class C on itself: we let true statements denote some function defined for every element of C , and false statements denote the function which is defined for no element of C . (This is clearly a wholly trivial move: it is clear that we can have *any* two distinct objects playing the role of the two truth values. However, note that in this setting, $\&$ denotes *functional composition*.) Then, we let different true statements denote *different* functions defined everywhere on C , and treat *any* such function as the truth value *the truth*. We can, for example, identify C with the universe and define $\|P(a)\|$ to be such function that for every $x \in C$, if $\|a\| \in \|P\|$, then $\|P(a)\|(x) = \|a\|$, and if $\|a\| \notin \|P\|$, then $\|P(a)\|(x)$ is undefined. $\|S\&S'\|(x)$ can then be defined as $\|S'\|(\|S\|(x))$ (where this is meant to be undefined where $\|S\|(x)$ is undefined). All of this is still trivial in the important sense of not tampering with the logical properties of the calculus. Now, however, we can easily provide an adequate semantics for statements containing \leftarrow : we can define $\|P(\leftarrow)\|$ to be such function that $\|P(\leftarrow)\|(x) = x$ if $x \in \|P\|$, and is undefined otherwise. And if we do this, the inferences (\leftarrow -I) and (\leftarrow -E) hold.

The final step is then to project the new semantics on terms. We can define the semantics e.g., in the following way (we can take S to be true if $x \in \|S\|(x)$ is defined for every element x of the universe; or else if it is defined for at least one element x of the universe - the difference affects only formulas with the “free occurrence” of \leftarrow , i.e. formulas in which \leftarrow does not follow an individual constant):

$\|a\|$ is a constant function defined everywhere on U (where a is an individual constant)

$\|\leftarrow\|$ is an identical function defined everywhere on U

$\|P(x)\|$ is a function such that $\|P(x)\|(y) = \|x\|(y)$ if $\|x\|(y) \in \|P\|$,

and is undefined if $\|x\|(y) \notin \|P\|$

$\|S\&S'\|$ is the composition of $\|S'\|$ and $\|S\|$ i.e. $\|S\&S'\|(x) = \|S'\|(\|S\|(x))$

In this way we turn denotations of sentences and terms into *context-change potentials*; and we create an *individuary* (a particularly tiny one, which can contain at most one individual).

¹² The enterprise is of course only a minor variation of the basic theme of dynamization of semantics as presented by Groenendijk & Stokhof (1991).

This creation results out of our attempt to account semantically for certain inferences - not to depict some mental or real machinery. In other words, the semantics based on this individuary has been employed as a *tool* of our account; not as a *picture*.

The introduction of the whole machinery of definite and indefinite singular terms, which leads to individuarities of more complex kinds, is now only a more complicated version of the same process. We have more inferences to account for: inferences like that of *an X Ps and Qs (a man walks and whistles)* from *an X Ps and the X Qs (a man walks and the man whistles)*; hence we need more 'slots' to store individuals, and we need labels to tell different slots apart. However, the individuary is again no more than a creature of our theory of drawing inferences.

The crux of this kind of semantic treatment is that the link between a singular noun phrase and a following pronoun (or between an indefinite noun phrase and a following definite noun phrase) is established by linking both phrases to the same inhabitant of the individuary. Saying that two noun phrases are *coreferential* becomes a short way of stating that their inferential roles are in a certain way interconnected (to say that within *A man walks and he whistles*, *a man* is coreferential with *he* is to say that the sentence implies *A man walks and whistles*), and talking about reference in turn becomes only a particularly illustrative way of rendering coreference. Reference is the relation between expressions and inhabitants of individuarities, but as individuarities and their inhabitants are mere tools of our accounting for inferences, the talk about reference is essentially parasitic upon the talk about inference - a *referent* is nothing more than an illustrious clamp holding certain inferentially related expressions together.

2.2 Capturing Inference as Reference

In general, we can see two expressions as inferentially connected if, informally stated, the inferences licensed by one of them are licensed also by the other. There are two levels of inferential relationships between expressions, the first level concerning (*material*) *implication* (licensing inferencing 'here and now'), and the second concerning *entailment* (licensing inferencing 'everywhere and always'). We say that a statement *S* *implies* a statement *S'* if *S'* is not false unless *S* is, i.e. if $S \rightarrow S'$; and we say that *S* *entails* *S'* if *S'* *cannot* be false if *S* is not; i.e. if $\Delta S \rightarrow S'$. Let us then call an expression *e* *weakly (inferentially) subordinated* to an expression *e'* iff for every atomic statement *S*, *S* implies $S[e'/e]$ (where $S[e'/e]$ is the statement which arises from *S* by replacing *e* by *e'*); and let us call *e* *strongly (inferentially) subordinated* to *e'* iff for every atomic statement *S*, *S* entails $S[e'/e]$. Finally, let us call *e* and *e'* *weakly (inferentially) equivalent* iff *e* is weakly subordinated to *e'* and *e'* is at the same time weakly subordinated to *e* (i.e. iff $S \leftrightarrow S[e'/e]$ for every statement *S*); and analogously for *strong (inferential) equivalence*.

With the help of this terminology, we can say something about the inferential behavior of names (nominal phrases). First, names are characterized by the fact that no name is subordinated to another name without being equivalent to it. This means that the inferential

structure of the domain of names is more or less trivial (in contrast to that of predicates which constitutes a Boolean algebra.) We inferentially characterize a name if we say with which other names it is weakly equivalent. If we say *coreferential* instead of *weakly equivalent*, then we can say that to inferentially characterize a name is to state with which other names it is coreferential; and as the relation of coreferentiality is clearly an equivalence relation which thereby decomposes the class of names into the corresponding equivalence classes, this is to specify the *coreferentiality class* to which the name belongs. And if we further say *what it refers to* instead of *which coreferentiality class it belongs to*, we can say that by inferentially characterizing a name we pinpoint what it refers to - thus explicating 'referent' as 'that which is shared by all coreferential expressions'.

Then we can distinguish several kinds of names. Some names, call them *proper* names, have the property that they are weakly equivalent (coreferential) if and only if they are strongly equivalent. Coreferentiality is thus a *standing* property for proper names, and the inferential behavior of a proper name is thus *exhaustively* characterized by specifying its (standing) coreferentiality class, i.e. by its (standing) referent.

Then there is another class of names - call them *descriptions*. Two descriptions can be coreferential without being strongly equivalent, and no description can ever be strongly equivalent to a proper name (although it can be coreferential with it). The relation of coreferentiality among descriptions is a fluctuating, contingent matter. Thus, the exhaustive inferential characterization of a description cannot consist simply in pointing out its (momentary) coreferentiality class; we must somehow say to which coreferentiality class it belongs when. This could be done by specifying its coreferentiality class relatively to the truth-valuation of sentences, i.e. to the possible world. Thus, a description can be inferentially characterized by being assigned a function from possible worlds to referents; as it is the case within montagovian frameworks.

Now there remains a third group of names, which can be called *pronouns*. No pronoun is coreferential with a proper name nor with a description; thus, we may be tempted to say that pronouns are 'in fact' not names and have no referents. However, pronouns are what can be called "strongly locally" equivalent with names - sentences containing pronouns are often strongly equivalent with sentences containing proper names in their place (*viz* ($\leftarrow I$) and ($\leftarrow E$)). One way to accommodate this is to make coreferentiality, and hence reference, somehow relative to the context - to associate pronouns with functions from contexts to referents. It then becomes vital to articulate the notion of context appropriately (as discussed in Section 1.4) and to embody this into a tractable compositional semantics.

Thus, we render the concept of reference, and referential links between expressions and occupants of 'slots' of individuarities, as a means to characterize inferential behavior. What, then, about the common temptation to see such 'slots' as depicting some places within the minds or brains of the speakers who draw the inferences? In a certain weak sense, this need not be incompatible with the vantage point advocated here: if we accept the inseparability of language and thought, we have to see *any* account of one's usage of language as *eo ipso* an

account of one's thinking. However, what I cannot find any substantiation for is seeing of the referential apparatus as the *depicting* of structures and processes going on within speakers' minds/brains (and thus holding that to assess a semantic theory we should observe what is going on within our heads).

Chomsky (1986, p. 45) writes: "One can speak of 'reference' and 'coreference' with some intelligibility if one postulates a domain of mental objects associated with formal entities of language by the relation with many of the properties of language, but all of this is internal to the theory of mental representations; it is a form of syntax." If we use the term *syntax* in such a way that we see inference as a syntactic matter (which is usual, although perhaps misleading), then the standpoint advocated here can be seen as consisting precisely in taking the theory of reference as "a form of syntax". However, Chomsky's pronouncement is problematic, it seems to me, because he speaks about "*mental* objects" and "*mental* representations" - which I find simply unwarrantable. I suggest that we replace the term "mental" in claims of this kind simply with the neutral term "semantic": perhaps some of the people speaking about *the mental* in this context really do *not* mean anything over and above *the semantic*.

2.3. Two Notions of Reference

So we have sketched how we can render the concept of reference as parasitic upon the concept of inference (and how we can see the individuary as a tool of a theoretician's account for inferences - not as a picture of a mental reality). However, it will be probably objected that the concept of reference is too substantial to be explained away in this simple way. Is reference not the very linkage between things and words which makes up our language in the first place? Is reference not the *cause* of inference?

It is often taken for granted that language is a higher stage of something which could be called "protolanguage" and which consists in direct reacting to things. First, a particularly naive version of the story might go, we have come to develop names of things which were relevant for us: seeing a tasty plant caused us to ejaculate one kind of shriek, seeing a horrible animal another; and when the shrieks became uniform enough from one case of seeing the plant or the animal to another, they habitualized themselves into *names* which *referred* to plants or to animals. Then, as our system of names slowly developed in richness and complexity, it became a *language*.

Seeing language in this way leads us to understand reference according to what Evans (1982) called the 'Photograph Model': as a photograph is a photograph of that which has brought it into being (by reflecting light rays into the lens of a camera), so a word is the name of that which has made the word come into being - which made us ejaculate the shriek which was the protoform of the name. The relation between a word and the thing which the word refers to is, according to this view, a matter of a 'causal chain' leading from the thing to the word. Moreover, it is this very relation which establishes the connection between words and things and which thus makes a system of shrieks and scribbles into *language*: thus, it is the relation of

reference on which all other concepts we use to describe language (meaning, truth, inference etc.) should be seen as resting - or, to use a word presently more fashionable, as *supervening*. This yields the program which has been outlined, e.g., by Field (1972): the program of reducing these notions to the causally understood concept of reference.

However, there is also a different view on language: according to this one, nothing deserves the name *language* unless it is a tool with which we can “play the game of giving and asking for reasons” (Brandom, 1994). This is to say, whatever the factual relation between a ‘protolanguage’ of the above described kind and a real language may be, language is *not* only a rich and complexly interrelated system of names. To think so is to fall with what Quine (1969) calls the *museum myth*: with the view that our words hook on things of the world as labels on exhibits of an immense museum. According to the story told by Brandom (but in fact going back to Kant), the distinctive feature of language is that it is capable of conveying *propositional content* - and hence that the constitutive characteristics of words is not that they name (refer to, represent, stand for) things, but rather that they can add up to propositionally contentful utterances. If we subscribe to this story, then we have to conclude that it is the relation of *inference*, not that of *reference*, which is the backbone of language on which everything else should be seen as supervening - for it is only *inferential articulation* which can confer propositional contents on sentences. As Brandom (*ibid.*, p. 84) puts it, “an association [of abstract objects with strings] amounts to specifically *semantic* interpretation just insofar as it serves to determine how those strings are correctly used”. The relation of reference is then seen as supervening on the relation of inference - which inevitably leads to a *deflationary* view of reference (in the sense of Horwich, 1990, Sec. 39).

In some of my recent writings (see esp. Peregrin, 1995) I have urged distinguishing between two views of language, which I have proposed to call the *nomenclatural* and the *structural* view, respectively (Brandom speaks about the *representational* and the *inferential* view to the same effect). The former view is based on seeing language as a nomenclature of some kind of things; the latter view sees language rather as a kind of toolbox. I have tried to indicate that it is the latter which can help us gain real insight into the workings of language. According to this view, the unequivocal aim of semantics is to account for the relation of inferability among statements - and any semantic value which semantic theory associates with an expression should be seen not as a depiction of a real thing on which the expression is claimed to (causally or otherwise) hook, but rather as a kind of hypostasis of the way the expression functions within inferences¹³. (This also implies that what concerns logic proof theory is in an important sense primary to model theory.) This view has several important consequences:

1. The way of semantics is essentially parallel to the way of logic - for logic is precisely the

¹³ Of course that if we want to see language generally as a matter of inferencing, then we have to construe the term *inference* broadly enough to comprise what Sellars (1974) calls *language entry transitions* (roughly inferences from situations to statements) and *language exit transitions* (inferences from statements to actions). However, these do not yield a nontrivial *theory* over and above trivialities like *We say ‘A rabbit runs’ iff a rabbit runs*.

account for inferability. The main difference is that semantics endeavours to see the inferential behavior of an expression materialized into its *meaning*, i.e. into an object associated with the expression.

2. Semantic formalisms should not be seen as *describing* platonistic entities or mental representations 'behind' expressions, but rather as *explicitly articulating* expressions' inferential properties. That is to say, a formula or a schema associated with a sentence should be seen as explicating the 'inferential potential' of the sentence, i.e. as summarizing what is implied by the sentence and what implies it; and the formula associated with a subsentential expression as that which adds up to the inferential potentials of the sentences in which it occurs.

Accordingly, I conclude that dynamic semantic frameworks can be understood in two different ways: as frameworks for capturing causal relations between words and things and as frameworks for capturing inferential relations between sentences. Here I want to suggest that it is the latter understanding to which we should adhere.

References

- Barwise, J., Cooper, R. (1981): 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* 4, 159-219.
- Brandom, R. (1994): *Making It Explicit*, Harvard UP, Cambridge (Mass.).
- Carnap, R. (1955): 'Meaning and Synonymy in Natural Languages', *Philosophical Studies* 7, 33-47 (later reprinted as an appendix to the author's book *Meaning and Necessity*).
- Chomsky, N. (1986): *Knowledge of Language*, Praeger, Westport.
- Egli, U. and K. von Heusinger (1995): *Choice Functions in Natural Language Semantics*, Universität Konstanz, Konstanz.
- Evans, G. (1982): *The Varieties of Reference*, Clarendon, Oxford.
- Field, H. (1972): 'Tarski's Theory of Truth', *Journal of Philosophy* 69, 347-75.
- Groenendijk, J. and M.Stokhof (1991): 'Dynamic Predicate Logic', *Linguistics and Philosophy* 14, 39-101.
- Groenendijk, J., M.Stokhof and F.Veltman (1996): 'Coreference and Modality', *The Handbook of Contemporary Semantic Theory* (ed. S.Lappin), Blackwell, Oxford.
- Heim, I. (1982): *The Semantics of Definite and Indefinite Noun Phrases* (dissertation), University of Massachusetts, Amherst.
- Heim, I. (1992): 'Presupposition Projection and the Semantics of Attitude Verbs', *Journal of Semantics* 9, 183-221.
- Hilbert, D. (1925): 'Über das Unendliche', *Mathematische Annalen* 95, pp. 161-190; translated as 'On the Infinite' in J. van Heijenoort (ed.): *From Frege to Gödel: A Source Book from Mathematical Logic*, Harvard University Press, Cambridge (Mass.), 1971, 367-392.
- Horwich, P. (1990): *Truth*, Blackwell, Oxford.

- Kamp, H. (1981): 'A Theory of Truth and Semantic Representation', in *Formal Methods in the Study of Language* (ed. by J.Groenendijk, T.Janssen and M.Stokhof), Mathematical Centre, Amsterdam; reprinted in *Truth, Interpretation and Information* (ed. by J.Groenendijk, M.Stokhof and T.Janssen), Foris, Dordrecht, 1984, 1-41.
- Lavine, S. (1994): *Understanding the Infinite*, Harvard University Press, Cambridge (Mass.), 1994.
- Meyer Viol, W.P.M. (1995): *Instantial Logic* (dissertation), Institute for Logic, Language, and Computation, Amsterdam.
- Montague, R. (1974): *Formal Philosophy: Selected Papers of R.Montague* (ed. by R.Thomason), Yale University Press, New Haven.
- Peregrin, J. (1995): *Doing Worlds with Words*, Kluwer, Dordrecht.
- Peregrin, J. and K. von Heusinger (1995): 'Dynamic Semantics with Choice Functions', in Egli and von Heusinger (1995), 41-67.
- Prawitz, D. (1965): *Natural Deduction*, Almqvist & Wiksell, Stockholm.
- Quine, W.V.O. (1969): *Ontological Relativity and Other Essays*, Columbia University Press, New York.
- Russell, B. (1905): 'On denoting', *Mind* 14, 479-493.
- Russell, B. and A.N.Whitehead (1913): *Principia Mathematica* I, Cambridge.
- Sellars, W. (1974): 'Meaning as Functional Classification', *Synthese* 27, 417-437.
- Strawson, P.F.(1950): 'On Referring', *Mind* 59, 320-344.