FORMAL REPRESENTATIONS OF SALIENCE IN DYNAMIC SEMANTICS

Simon Varey

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy in the College of Arts and Sciences

Chapel Hill
2019

Approved by:

Gillian Russell

Thomas Hofweber

Ram Neta

ABSTRACT

Simon Varey: Formal Representations of Salience in Dynamic Semantics
(Under the direction of Gillian Russell)

Lewis' (1973; 1979) salience analysis of definite description reference suggests that an utterance of 'the F' refers to the most salient F, according to a contextual salience ranking. This analysis was created to account for referring improper definite descriptions. As shown by von Heusinger (2004), the salience analysis can also be used to explain anaphoric definite descriptions. Any formalization of the salience analysis will require a method for formally representing salience, such as Lewis' order-theoretic method or von Heusinger's choice-theoretic method. In this thesis, I will argue that the order-theoretic method explains certain features of salience that the choice-theoretic method does not, and therefore the order-theoretic method should be favored. I will also show how von Heusinger's Dynamic Semantics with Choice Functions can be adapted to utilize the order-theoretic method, thus providing a formal modeling of the context change described by the salience explanation of anaphoric definite descriptions.

ACKNOWLEDGEMENTS

I would like to acknowledge the vital assistance and guidance provided by the chair of my committee Gillian Russell, and the other members of my committee Thomas Hofweber and Ram Neta. I would also like to thank Dean Pettit and the audience at my work-in-progress talk at the University of North Carolina for their comments. Finally, I would like to thank my parents Doug and Sue, my sister Katrina, and my partner Erin for their love and the support they provide me in all my academic pursuits.

TABLE OF CONTENTS

1 The Salience Approach to Definite Descriptions

1.1 The Problem of Improper Definite Description Utterances

A *theory of reference* for definite descriptions must tell us, for every possible utterance of a definite description, what that utterance refers to.[1] The reason that the reference of definite descriptions must be predicted relative to utterances of definite descriptions is that different utterances of the same definite description can have different referents. Another way of putting the same point is that definite description reference is determined relative to *contexts*, where contexts contain all the contextual factors that can affect definite description reference. Following Kaplan (1980), I will represent a context c as a tuple containing a possible world w_c, and various parameters defined in that world. The domain of w_c can also be called *the domain of the context* c, and it represents everything that exists in c.

The constituent phrase 'F' of a definite description 'the F' is called the *restrictor* of that definite description. We can call the set of entities described by the restrictor of a definite description 'the F' within a given context c the *restrictor set* of 'the F' in c. I will represent the denotation of an utterance of a phrase P as $\|P\|^c$,[2] so the restrictor set of 'the F' in c will be $\|F\|^c$. Every utterance of a definite description therefore has a restrictor set, being the restrictor set of that definite description in the context of the utterance. Using the notion of a restrictor set, we can divide utterances of definite descriptions into three kinds: *empty* definite description utterances (whose restrictor sets are empty), *proper* definite description utterances (whose restrictor sets are singletons), and *improper* definite description utterances (whose restrictor sets have multiple members).

With this three-way division in mind, consider the following principles of definite description reference:

(UE) Empty definite description utterances have no referents.

(UP) Proper definite description utterances refer to the single member of their restrictor sets.[3]

(UI) Improper definite description utterances have no referents.

[1] For the purposes of this thesis, I will put aside Russell's (1905) claim that definite descriptions are quantificational expressions. I will also put aside plural definite descriptions (i.e. phrases of the form 'the Fs').

[2] The denotation of an utterance of a predicate will be its extension, and the denotation of a term will be its referent.

[3] Note that I am only considering what Donnellan (1966) called the *attributive* use of definite descriptions in this thesis.

Together, these principles can be referred to as the *uniqueness analysis* of definite description reference, as they suggest that definite description utterances only refer when their restrictors uniquely describe a particular entity within the utterance's context of use. This view of definite description reference was first clearly stated in Frege's 1892 paper "On Sense and Reference", and has proven to be one of the most popular theories of definite description reference.

The uniqueness theory does seem to make the correct predictions when it comes to empty and proper definite description utterances. However, it seems to make entirely the wrong predictions when it comes to improper definite description utterances. Certainly, some improper definite description utterances do not seem to have any referent. If I utter the definite description 'the table' in a room with multiple tables (as far as I can tell), none of which have either been specifically mentioned before or are drawing attention to themselves in any way, then it does seem as if my utterance has no reference. However, many improper definite description utterances do seem to have referents. If I utter the definite description 'the cat' in a room with a single cat Max (as far as I'm aware), then it seems entirely possible (indeed likely) that I am referring to Max. However, so long as there is more than one cat in the universe, the uniqueness analysis will not allow for this possibility. Thus, we can see that improper definite description utterances which *do* refer create a problem for the uniqueness analysis. I will call this problem *the problem of improper definite description utterances.*

A solution to this problem would involve replacing the uniqueness analysis with a theory of the reference of definite descriptions which makes the same predictions in the case of empty and proper definite description utterances, but which makes a different prediction in the case of improper definite description utterances. Several such solutions have been suggested, but one of the most successful and powerful such solutions is Lewis' (1973; 1979) *salience analysis* of definite description reference.

1.2 The Salience Analysis of Definite Description Reference

Consider this example, from Lewis (1979, p. 348, my emphasis):

> In the room is a cat, Bruce, who has been making himself very salient by dashing madly about. He is the only cat in the room, or in sight, or in earshot. I start to speak to you: '*The cat* is in the carton. The cat will never meet our other cat, because our other cat lives in New Zealand. Our New Zealand cat lives with the Cresswells. And there he'll stay, because Miriam would be sad if *the cat* went away.'

Intuitively, the first utterance of 'the cat' here refers to Bruce, while the second utterance of 'the cat' refers to a different cat who lives in New Zealand, who Lewis calls Albert. Note, however, that this would be impossible on the uniqueness analysis. On that analysis, the first utterance of 'the cat' could only refer to Bruce if he were the only cat in the context, but if that were the case the second utterance of 'the cat' would have to refer to Bruce as well. As such, Lewis (1979, p. 348) concludes "it is not true that a definite description "the F" denotes x if and only if x is the one and only F in existence."

Why is it that, intuitively, the first utterance of 'the cat' refers to Bruce and the second to Albert, when (presumably) millions of the cats exist in the context? The reason, Lewis suggests, has to do with how *salient* those cats are to the conversation. Lewis (1979, p. 348) notes that "at first, "the cat" denotes Bruce, he being the most salient cat for reasons having nothing to do with the course of conversation." These reasons are the fact that he is 'dashing madly about', and because 'he is the only cat in the room, or in sight, or in earshot'. Bruce's salience means that, when Lewis wants to talk about Albert, he has to use definite descriptions that *don't* describe to Bruce, such as ""our other cat" or "our New Zealand cat"" (Lewis, 1979, p. 348). However, by talking about Albert, Lewis "raise[s] Albert's salience by conversational means" which means that, by the second utterance, he is "in a position to say "the cat" and thereby denote not Bruce but rather the newly-most-salient cat Albert" (Lewis, 1979, p. 348-9).

Lewis (1979, p. 348) then presents the following analysis of definite description reference: ""the F" denotes x if and only if x is the most salient F in the domain of discourse, according to some contextually determined salience ranking." Note that this analysis does not tell us "what happens if two F's are tied for maximum salience, or if no F is at all salient" as Lewis wants to "pass over [these] complications" (1979, p. 348). I will however suggest that in these cases the salience analysis should hold that utterances of 'the F' do not refer to anything. I will defend this position below. Using the terminology introduced above, the *salience analysis of definite description reference* can then be encapsulated by the following principle:

(S) Definite description utterances refer to the uniquely maximally salient member of their restrictor sets, if there is such a member, and nothing otherwise.[4]

What is the notion of salience that is being utilized in this principle? That will turn out to be a difficult question to answer. In the next two subsections, where I will outline a method to formally represent salience, and note some formal constraints that must be placed on this method if it is to allow the salience analysis to be successful, I will take salience to be a primitive notion. In the final subsection of this section, I will elaborate what I take this notion of salience to be.

[4]For simplicity, I will refer to the uniquely maximally salient member of a set as the 'maximally salient' or 'most salient' member of the set.

1.3 The Ordering Method of Salience Representation

If the salience analysis is to be formalized as theory of reference, it will require some *method* for formally representing these "contextually determined salience ranking[s]" (Lewis, 1979, p. 348). Lewis (1973, p. 112-3) considers a *system of spheres method*, where salience in general is represented by "a system of spheres $ based on comparative salience of things."

When it comes to the counterfactuals and "comparative similarity of possible worlds" (p. 8), Lewis (1973) defines $ as "an assignment to each possible world i of a set $\$_i$ of sets of possible worlds" (p. 13-14). $ is then a "*system of spheres*, and members of each $\$_i$ are called *spheres* around i, if and only if, for each world i, the following conditions hold ...

(1) $\$_i$ is *nested*; that is, whenever S and T belong to $\$_i$, either S is included in T or T is included in S.

(2) $\$_i$ is *closed under unions*; that is, whenever S is a subset of $\$_i$ and $\bigcup S$ is the set of all worlds j such that j belongs to some member of S, $\bigcup S$ belongs to $\$_i$.

(3) $\$_i$ is *closed under (nonempty) intersections*; that is, whenever S is a nonempty subset of $\$_i$ and $\bigcap S$ is the set of all worlds j such that j belongs to every member of S, $\bigcap S$ belongs to $\$_i$" (p. 14).[5]

Lewis (1973, p. 113) adapts this notion to represent salience, such that "when i is a thing with a point of view ... a sphere around i is to be any set of things in the ken of i such that all those in the set are more salient to i than any of those outside." We can therefore see a nested set of spheres $\$_i$ as representing a ranking of how salient entities are to an agent i. Entities in the smallest sphere are ranked highest in salience, while entities in the second smallest sphere (but not in the smallest sphere) are ranked second, and so on.

Lewis' approach, whereby spheres can contain only things in the 'ken' of an agent, leads to the consequence that all those things the agent doesn't know about are not anywhere in their salience ranking. Lewis (1973) calls a system of spheres *universal* when its largest sphere contains everything in the relevant domain (p. 16) and claims that "the system $ [representing salience] is not universal ... since for most

[5]In the similarity case, Lewis (1973, p. 14) considers *centered* systems of spheres which obey the following condition: "$\$_i$ is *centered on* i; that is, the set $\{i\}$ having i as its only member belongs to $\$_i$." In the salience case which I will discuss below, this condition would be equivalent to every agent being most salient to themselves. As Lewis (1973, p. 113) notes, "we are not all such egoists; for at least some normal i, there is something even more salient to i than he is to himself." As such, systems of spheres are not centered in the salience case.

things i there is something outside the ken of i" (p. 113). But this seems like the wrong result. Those things are somewhere in the ranking, namely right at the bottom. Thus, I will hold that everything in a context, not just those things an agent knows about, are in at least one sphere, with the things an agent doesn't know about in only the largest sphere.[6]

Lewis understands salience relative to agents, but we can instead understand it relative to contexts, and define $\$_c$ as the salience ranking of a context c,[7] which would mean c is the ordered pair $\langle w_c, \$_c \rangle$.[8] To formalize the salience analysis using this method of salience representation, I first need to define a function $\max_\$$, such that for any salience ranking $\$_c$ and subset X of the domain of c, $\max_\$(\$_c, X)$ is the smallest sphere in $\$_c$ with a non-empty intersection with X.[9] (S) can then be defined formally as:

$$\text{(SS)} \qquad \|\text{the } F\|^c = \begin{cases} x & \text{if } \|F\|^c \bigcap \max_\$(\$_c, \|F\|^c) = \{x\} \\ \text{nothing} & \text{otherwise} \end{cases}$$

Lewis (1973, p. 48) also notes, in the counterfactuals case, that "our system of spheres is nothing but a convenient device for carrying information about the comparative similarity of worlds. We could do away with the spheres, and give the truth conditions for counterfactuals directly in terms of comparative similarity of worlds." Similarly, in the definite descriptions case, we could do away with the spheres, and give the reference conditions for definite descriptions directly in terms of comparative salience.

Thus, let's consider an *ordering method* of representing salience. Using this method, salience is represented by a binary relation $\geq$ between entities in a context of use, such that $x \geq y$ holds if and only if an entity x is no less salient than an entity y.[10] Each context c will have its own *salience ordering* $\geq_c$

[6] Lewis (1973, p. 14) notes that, in the counterfactuals case, "we may ... count the set of *all* worlds as one of the spheres around i" given "it vacuously meets all the conditions above". As such, this move does not violate any of the constraints of (1)-(3). Furthermore, Lewis (1973, p. 16) "regard[s] the worlds that the spheres around i do not reach—those that lie outside $\bigcup \$_i$ —as all being equally similar to i, and as less similar to i than any world where the spheres do reach," which is just the same as he would regard those worlds in the largest sphere in a universal system. As such, it may seem like there is no difference between treating these worlds, or entities in the definite descriptions case, as being in the largest sphere versus not being in the system at all. There is, however, a difference which I will demonstrate in the next subsection.

[7] I will discuss the relationship between these two notions in the next subsection.

[8] Of course, contexts in the final account of language will have many more parameters than just salience, but as salience is all that is necessary for determining the reference of definite descriptions, this simplification will be harmless here.

[9] If every sphere in $\$_c$ has an empty intersection with X then $\max_\$(\$_c, X)$ can be defined as the empty set. Note that, given $\$_c$ is universal, this will only arise where $X = \emptyset$. I will discuss the important of this point in the next subsection.

[10] This ordering can be used to define a strict salience ordering $>$ such that $x > y$ holds if and only if x is more salient than y. The definition is as follows: $x > y$ if and only if $y \not\geq x$.

defined between the entities in that context, so that $c = \langle w_c, \geq_c \rangle$. I will call the subset of members of a set X which are maximal relative to $\geq$ $\max(\geq, X)$.[11] (S) can then be put formally as:

$$\text{(SO)} \qquad \|\text{the } F\|^c = \begin{cases} x & \text{if } \max(\geq_c, \|F\|^c) = \{x\} \\ \text{nothing} & \text{otherwise} \end{cases}$$

Lewis (1973, p. 48) places two conditions on this binary relation in the similarity case, which also apply to the salience case, namely *transitivity* and *totality* (a.k.a. *strong connection*):

(Totality) For all entities x and y in a context c, either $x \geq_c y$ or $y \geq_c x$ (or both).

(Transitivity) For all entities x, y and z in a context c, if $x \geq_c y$ and $y \geq_c z$ then $x \geq_c z$.

Relations with both these properties are called *weak orderings* (a.k.a. *total preorderings*). Totality implies that weak orderings are also *reflexive* (i.e. for all entities x in a context c, $x \geq_c x$). Note that weak orderings are *not* in general *anti-symmetric* (i.e. there can exist two entities x and y in a context c such that $x \geq_c y$ and $y \geq_c x$ but it is not the case that x and y are identical). I will explain why salience orderings have the properties of weak orderings in the next subsection.

Lewis (1973, p. 49) notes that this "formulation is exactly equivalent to the original formulation by means of spheres." In the similarity case, this means that "we can put systems of spheres in one-to-one correspondence with comparative similarity systems [i.e. orderings], in such a way that the corresponding systems agree on the truth value at every world of every counterfactual" (Lewis, 1973, p. 49) . The same is true in the salience case. Let's say that two representations A and B of the salience of the entities in a possible world w_c are equivalent where, for every definite description 'the F', 'the F' has the same referent (or lack thereof) in $\langle w_c, A \rangle$ as it does in $\langle w_c, B \rangle$. We can then show that, for every salience sphere system $\$_i$, we can define an equivalent salience ordering $\geq_i$, and vice versa.[12] We can then say that two *methods* of salience representation are equivalent if and only if there is a one-to-one relation between representations in one method and *equivalent* representations in the other. Given there is such a relation between $\$_i$ and $\geq_i$, the system of spheres method and the order-theoretic method are equivalent.

[11] Formally, $\max(\geq, X) = \{x \in X : \forall y \in X(x \geq y)\}$.

[12] The proof follows the one Lewis (1973, p. 49) provides.

As such, the choice between the system of spheres method and the order-theoretic method is for the most part a technical one. Lewis (1973, p. 113) himself does not appear to see much difference between them as he uses the two methods interchangeably: "assuming that comparative salience orderings have the definitive properties of weak orderings, then these sets do indeed comprise a system of spheres." For the rest of this paper, I will treat Lewis as if he used only the ordering method, and will not discuss the system of spheres method. The main reason why I will do this is because it will be easier to think in terms of the two, relatively simple conditions of weak orderings rather than the three, more complicated conditions of systems of spheres.

1.4 The Adequacy of the Salience Analysis

For the salience analysis to provide a solution to the problem of improper definite description utterances, it must provide an adequate theory of the reference of definite descriptions. In other words, it must predict the correct referents in the cases of empty and proper definite description utterances, like the uniqueness analysis, while also making the correct predictions in the case of improper definite description utterances, unlike the uniqueness analysis. Let's consider the cases of empty and proper definite description utterances first. To make the correct predictions in these cases, the salience analysis must entail that (UE) and (UI) are true. I will now demonstrate how these claims are in fact entailed by the salience analysis.

Let's consider (UE) first: Empty definite description utterances have no referents. (S) entails that a definite description utterance will have a referent only if the utterance's restrictor set has a maximally salient member. Of course, if a restrictor set has *no* members, then *a fortiori* it will not have a maximally salient member. As such, (S) entails (UE), which means that the salience analysis makes the correct predictions in the case of empty definite description utterances. This is half of my justification for the claim that utterances of 'the F' do not refer if there is no most salient F. Note this this explanation does not make use of any of the special properties of weak orderings, and therefore would hold no matter what kind of orderings salience orderings were.

Let's now consider (UP): Proper definite description utterances refer to the single member of their restrictor sets. (S) entails that a definite description utterance will refer to the most salient member of its restrictor set. But if the definite description utterance is proper, then its restrictor set *only has* one member, which must then be the most salient member. (S) therefore entails (UP), which means that the salience analysis makes the correct predictions in the case of proper definite description utterances. As Krahmer (1998, p. 211) puts it, "if there is only one object with the right properties, it has to be the most

salient one with these properties." To put the point formally, (SO) entails that 'the F' in c refers to $\max(\geq_c, \|F\|^c)$ where $\max(\geq_c, \|F\|^c)$ is a singleton. If 'the F' in c is proper, then $\|F\|^c = \{x\}$ for some entity x. Given that $\geq_c$ is reflexive, we can prove that $\max(\geq_c, \{x\}) = \{x\}$[13] and therefore that 'the F' in c refers to x.

One thing to note, when it comes to the referents of proper definite description utterances, is that it is not necessary for an entity to be salient at all for it to be referred to by a proper definite description utterance. A current utterance of 'the largest city in the United States' refers to New York, irrespective of whether New York has been previously mentioned in the conversation, or is otherwise relevant to it. This fact is why, contrary to Lewis, salience *does* need to be universal. If New York was not within the salience ordering, then we could not even say it was equally salient to itself, and if that were the case then the salience analysis would not predict that an utterance of 'the largest city in the United States' would refer to New York.[14]

That just leaves improper definite description utterances. The first thing to note is that (UI) is not true on the salience analysis, and therefore the salience analysis does not make the same error that the uniqueness analysis makes. We can see this by considering a simple counterexample. Suppose we are in a context C, where there are only two cats: Max and Zac. In C, the restrictor set of an utterance of 'the cat' would be {Max, Zac}. Given there is no unique cat in C, the uniqueness analysis would predict that an utterance of 'the cat' in C would not refer to anything, in accordance with (UI). Let's then suppose that Max is more salient than Zac in C, which means that Max is the maximally salient cat in C. Thus, under the salience analysis, an utterance of 'the cat' in C would refer to Max. This means that C is a counterexample to (UI) under the salience analysis. However, to note that the salience analysis does not make the same error that the uniqueness analysis does when it comes to improper definite description utterances is not yet to say that it provides the correct analysis. To determine whether the salience analysis does provide the correct answer will require a more precise notion of salience, which I will present in the next subsection.

[13] Proof: for all y in X, $y \in \max(\geq, X)$ if and only if $y \in X$ and, for all z in X, $y \geq z$, by the definition of the max function. As such, $x \in \max(\geq, \{x\})$ if and only if, $x \in \{x\}$ (which is true) and $x \geq x$. Given that $\geq_c$ is reflexive, $x \geq_c x$ holds, which means $x \in \max(\geq_c, \{x\})$. Furthermore, nothing other than x can be a member of $\max(\geq_c, \{x\})$, so $\max(\geq_c, \{x\}) = \{x\}$. This result is a reason why salience orderings have to be reflexive.

[14] The source of Lewis' error here is, I think, a disanalogy between the counterfactual and definite description cases which Lewis does not seem to have appreciated. He says, of worlds outside a similarity system $\$_i$, "we will see that any such world will be left out of consideration in determining whether a counterfactual is true at i. It is as if, from the point of view of i, these remotest worlds were not possible worlds at all" (Lewis, 1973, p. 16). Note, however, that no entity can be left out of consideration when it comes to definite description reference, as even the least salient entity may be referred to by a proper definite description utterance.

1.5 The Notion of Salience

Lewis (1973, p. 113) says the following about salience: "When i is a thing with a point of view ... then some things are more salient than others from the point of view of i. They loom larger in his mental life; they are more important to him; they come more readily to the center of his attention." This, of course, would not serve as a very useful definition of salience, as 'importance' and 'attention' are largely synonymous with salience.[15] Other terms we might use to define salience, such as 'relevance' and 'interest', have similar issues. However, I think we can use this discussion to ostend a particular intuitive notion. It seems clear enough that at a given time we may have certain entities in mind, while other entities escape our interest. What the salience analysis claims is that it is in virtue of this mental ranking that a particular F is selected as the referent of an improper utterance of 'the F'.

Two things should be noted at this point. Firstly, I should discuss the sense in which I see salience as mental. The idea is that the mental attitudes of an agent has toward an entity determine how salient that entity is to them. Note that this does not mean that an entity which an agent is unaware of cannot be anywhere on an agent's salience ranking, as Lewis seems to have thought. However, it does mean that all such entities must be at the *same* ranking, given that agents must have all the same attitudes to entities they are unaware of. The salience rankings I have discussed so far obey this condition, as they all put such entities at the bottom of the ranking. Note that this also means that salience is a subjective notion. An asteroid heading towards the Earth might, in some sense, be highly important for an agent, but if the agent doesn't know it exists then it won't be salient for them. This has the consequence that an agent cannot use the definite description 'the asteroid' to refer to such an asteroid, even if it is the most important (in an objective sense) asteroid to them. This seems like the right result.

Secondly, it might be unclear how a notion of salience to an agent is related to the notion of salience relative to a context I discussed above. I can see two ways in which these notions could be related, depending on how we think utterances become meaningful. If we take the meaning of an utterance to be what the speaker of the utterance intends it to be, then presumably what a speaker intends an utterance of a definite description 'the F' to refer to is what *they* take to be the uniquely most salient F. Given this is the case, if we take the meaning of an utterance to be what the speaker of the utterance intends it to be, then we can identify the salience ranking of a context with the salience ranking of the speaker of that context.[16]

[15] This is not to suggest that Lewis intended it to be a definition.

[16] Note that this view would require contexts to have an additional 'speaker' parameter. Kaplan's (1980) original notion of a context of use includes such a parameter.

However, if we suppose that an utterance is meaningful only where it is understood in the same way by the speaker and the audience, then the salience ranking of the context will need to be different. To see why, consider the following: **(Example 1)** 'Lulu and Annie are in a room together. They both believe that their cat Zac is in the next room. Lulu notices their other cat Max run into the room, while Annie falsely believes Max is still outside.' In this case, it seems likely that Lulu and Annie's salience rankings are different.[17] Zac is the maximally salient cat for Annie, while Max is the maximally salient cat for Lulu. If one thought that only the speaker's intentions were relevant when it comes determining reference, then an utterance of 'the cat' by Lulu in this context would unproblematically refer to Max. Yet Annie would (seemingly justifiably) interpret Lulu's utterance as referring to Zac. If one considered this divergence of interpretation to sufficient to render Lulu's reference to Zac problematic, then one would need a different notion of the salience of a context.

One possibility along these lines is to utilize Stalnaker's (1978, p. 84) notion of a *conversational common ground*. For Stalnaker (1978, p. 84), "the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation" and the set of all such "presuppositions are what is taken by the speaker to be the COMMON GROUND of the participants in the conversation."[18] Above, I suggested that the set of attitudes an agent has towards the entities they are aware of determines their (private) salience ranking. In the same way, we can see the set of such attitudes presupposed by an agent within a context as also determining a salience ranking, which we can call their *presupposed salience ranking*. An agent's private and presupposed salience ranking can come apart. **(1)** provides an example of this: if Lulu knows that Annie doesn't know that Max came inside, then Lulu will not presuppose this fact (although she still believes it). Thus, on her private salience ranking Max is the uniquely most salient cat, while on her presupposed salience ranking Zac is.

Stalnaker (1978, p. 85) "define[s] a NONDEFECTIVE CONTEXT as one in which the presuppositions of the various participants in the conversation are all the same." In such a context, every participant will have the *same* presupposed salience ranking. *This* ranking can then be considered as the salience ranking of the context. **(1)** could be an example of a nondefective context. In this example, Zac would be the maximally salient cat according to the context's salience ranking, and therefore he would be the referent of Lulu's utterance. The question of what definite descriptions refer to (if anything) in a *defective* context is a

[17] I will provide further explanation of this fact below.

[18] Note that, for Stalnaker (1978, p. 84) "the propositions presupposed in the intended sense need not really be common or mutual knowledge; the speaker need not even believe them. He may presuppose any proposition that he finds it convenient to assume for the purpose of the conversation, provided he is prepared to assume that his audience will assume it along with him."

difficult one, which I will put aside here.[19] While I don't think that my major points in this paper are affected by the decision between these two ways of understanding the salience of a context, I will use the second understanding for the rest of this paper. This understanding seems better firstly because it seems to get the correct result in **(1)** (unlike the first understanding) and secondly because the first understanding problematically seems to suggest reference of definite description utterances is a private, not a public, matter.

A completely successful analysis of definite descriptions should be able tell us what any utterance of a definite description refers to, perhaps relative to certain contextual factors of the utterance. The salience analysis does tell us what any utterance of a definite description refers to *relative to the salience ranking of the context*. However, I have not yet said how to determine what the salience ranking of a context is. As such, the analysis cannot be considered completely successful, as it does not yet provide any predictions regarding the reference of improper definite description utterances.[20]

As salience is not a primitive notion, a reductive theory of salience, which would predict the salience ranking of any person at any time, should be possible. I believe that such a theory is in principle possible. However, creating it would require extensive, empirical psychological research, and it would no doubt be highly faceted, reflecting not only the person's current mental state but also their history, preferences and society. I will not attempt to present such a theory here. One might then fear that my analysis is empty, as it fails to provide a definite reference prediction in the problem cases of improper definite description utterance. Relatedly, one might fear that the theory is ad hoc, as I could just claim (in an unprincipled fashion) that the salience ranking is whatever I need it to be in a given context to get the result I need.

To avoid these worries, let me explain in more detail how I intend to invoke salience in my explanations. While I will not present a complete theory of salience, I do intend to suggest several broad principles for determining salience in a given context. I will then claim that these principles are sufficient for the explanations I want to provide. As such, even though I will not provide a full theory of salience, I will provide enough of a theory for my purposes. Furthermore, the principles I will invoke will be independently justified, thus showing how the analysis is not ad hoc.

[19] Stalnaker (1978, p. 85) allows for contexts that are "CLOSE ENOUGH to being nondefective if the divergences do not affect the issues that actually arise in the course of the conversation." In practice, very few contexts will be nondefective in the sense of participants presupposing the same salience for *every* entity in the context. However, so long as the participants' disagreement does not affect the reference of any definite description actually uttered in the context, we can treat the context as if it were nondefective.

[20] As we saw in the previous subsection, the salience analysis predicts the reference (or rather the lack thereof) of empty definite description utterances independently of any particular salience ordering, and the reference of proper definite descriptions with only the proviso that all salience orderings are reflexive.

The approach I take to salience is similar to the approaches taken in several other fields which invoke similar rankings in their explanations. One example is Lewis' (1973) use of similarity orderings of possible worlds in his analysis of counterfactual conditionals. Lewis does not provide an exact theory of the similarity of worlds, but he does offer several general, independently justified principles. We can see a similar form of explanation in the invocation of preference orderings in decision theory. We do not have a complete psychological theory which would predict any person's preferences at any time. However, decision theorists have proposed some very general (and independently justified) principles regarding preference orderings, and merely from these general principles decision theorists have been able to explain a lot about decision-making. One of the general principles that all three explanations invoke is that their respective rankings are all weak orderings.[21] Parallel to these explanations, by proposing some very general (independently justified) principles regarding salience orderings, I hope to explain a lot about definite description reference.

So what are these principles? We have already seen two: totality and transitivity. Totality entails that every entity is comparable in terms of salience. This principle seems plausible given our intuitive understanding of salience. The exceptions to this principle would be pairs of entities such that one could not discern which of the two seemed more salient. This might be the case if one's attitudes towards the entities were too impoverished to be able to discern between them in terms of the factors relevant to salience. However, I think that such cases are better understood as cases where it is a *vague* matter as to which of the entities is more salient (or whether they are equally salient), rather than cases where the entities are *incomparable* in terms of salience. I will suggest this because, when confronted with these cases like these, it seems that we adopt the same procedures of *precisification* that we do in other cases of vague properties. I will suggest that whatever approach one takes to vagueness in general will apply to vagueness in salience as well.

Transitivity entails that there are no cycles of salience. This principle also seems very plausible given our intuitive understanding of salience. If an entity x is more present to mind than an entity y, and y is more present to mind than an entity z, then I cannot conceive of z being more, or even equally, present to mind than x.

[21] One might wonder why similarity, preference and salience orderings are all weak orderings. Here is a very speculative proposal: we might suppose that similarity, preference and salience are not just orderings, but also *metrics*. If we suppose that every entity has a salience *value* (for a person at a time), and then we order those entities by their salience values, we will end up with a weak ordering. The same would be true if we ordered worlds by their similarity value or results by their preference value. There are two reasons why I am only speculating that salience is a metric in this thesis: a) there is no necessity for me to introduce a salience value, because only comparative (and not absolute) salience is relevant for the salience analysis, and b) the evidence that I have for salience (i.e. the reference of definite description utterances) is insufficient to reconstruct anything other than comparative salience.

There are two more principles I want to discuss. The first is this: mentioning or pointing to an entity raises its salience. This principle will be essential in my discussion of salience and anaphora in the next section. As such, I will elaborate, and justify, this principle in the next section. The second principle is this: *ceteris paribus*, entities believed to be closer to the conversational participants are more salient than those that are far away. This principle should be primarily seen as a heuristic, as it applies only when no other principle does. Furthermore, distance here should not be understood literally in terms of millimeters, but rather in terms of a more human system of distance. What I have in mind here is a system where (for a person in a building) entities believed to be in the same room are more salient than those believed to be in the building outside the room, those believed to be in the building are more salient than those believed to be outside, and so on. Note that this heuristic is compatible with the principles I mentioned above. It is compatible with totality, in so far as everything is believed to be *some* distance from a person, and with transitivity, in so far as the 'further away from' relation is transitive.

Note that this distance heuristic would explain the reference of the two cases of improper definite description utterances I mentioned in the Subsection 1.1. The first was this: 'If I utter the definite description 'the table' in a room with multiple tables (as far as I can tell), none of which either have been specifically mentioned before or are drawing attention to themselves in any way, then it does seem as if my utterance has no reference.' In this case, the distance heuristic would entail that all the tables in the room have the same salience (and more salience than any table in the rest of the world). This would mean that no table was uniquely maximally salient in this context, and as such under the salience analysis an utterance of 'the table' in this context would have no referent. Thus, the salience analysis gets the correct result in this case (as, incidentally, does the uniqueness analysis, albeit for the wrong reason). This is the other half of my justification for the claim that utterances of 'the F' do not refer if there is no most salient F.

The second was this: 'If I utter the definite description 'the cat' in a room with a single cat Max (as far as I'm aware), then it seems that I am referring to Max.' In this case, the distance heuristic would entail that Max is the maximally salient cat (even if other cats exist in the context), which means that the salience analysis would predict that an utterance of 'the cat' would refer to Max. Thus, the salience analysis gets the correct result in this case too (unlike the uniqueness analysis).

The distance heuristic also explains **(1)**. As far as Annie is concerned, Zac is the closest cat, so he is the maximally salient cat to her. Lulu thinks that Max is the closest cat, so he is the maximally salient cat on her private ranking, but she knows that Annie doesn't think Max is the closest cat, so she presupposes

Zac is closer, so he is the maximally salient cat on her presupposed ranking. Thus Zac is the most salient cat in the context, and an utterance of 'the cat' will refer to him.

Note that it is important that the distance heuristic operates in terms of *believed* distance, rather than actual distance. This is important because salience operates as a subjective notion. Consider the following example where these two notions come apart: Suppose, in **(1)**, that neither Lulu nor Annie knew that Max had come back inside. If actual distance was the pertinent property, then Max would be most salient in this content, while if believed distance were pertinent, then Zac would be most salient. As an utterance of 'the cat' in such a context would seem to refer to Zac, then the decision to focus on believed distance seems to be the correct one.

More generally, I will suggest that the problem of improper definite description utterances can be solved by the salience analysis along the lines of the specific cases I mentioned above. The few examples I have provided do not, of course, provide definitive proof that the salience analysis provides the correct prediction for every improper definite description utterance, but I hope that they illustrate and motivate the general approach. One question to ask is whether the principles I have mentioned are the only principles which govern salience. The answer to this, I think, is surely no. I am sure that further investigation would discover further salience properties of greater or lesser generality. This claim may open me up to further charges of ad hocery. It would seem that any time a counterexample to my analysis was raised, I could just introduce a new principle to explain it away. To reject this charge, I will note two further features of the notion of salience I am utilizing. Firstly, note that salience is an intuitive notion on my view, and not a stipulative notion. As such, whatever principles are introduced must be justified in terms of our intuitive notion of salience. Secondly, any new principle must be consistent with the principles that have already been introduced. As such, a counterexample which contradicts one of the stated principles cannot be explained away by introducing a new principle, and therefore will invalidate the theory.[22] The exception to this is the distance heuristic, which only holds *ceteris paribus*.

[22]This is separate from a counterexample which is neither predicted by nor contradicts one of the stated principles. Perhaps these cases are better understood as lacunae, rather than counterexamples. Regardless, such lacunae are problematic for the analysis, and the possibility that they could be merely resolved through postulation of further principles does threaten to make the analysis ad hoc.

2 The Dynamic Approach to Anaphora

2.1 Anaphoric Definite Descriptions and Salience

An utterance of a term is an *anaphor* when it refers to an entity x in virtue of an earlier term utterance referring to x. The earlier term utterance is called the *antecedent*, and together the two utterances possess an *anaphoric link*. Definite description utterances can be anaphora. Consider a slightly altered version of the example from Lewis (1979, p. 348) I discussed in Section 1. **(Example 2)** Suppose David began a conversation by uttering '*Albert* lives with the Cresswells, and there *he*'ll stay, because Miriam would be sad if *the cat* went away' in a context where no cat was close by or otherwise drawing attention to themselves in any way.[23] Albert is David's cat who lives with the Cresswell's. David's utterance of 'the cat' in this example seems to refer to Albert, and (given the irrelevance of any cats in the context prior to this utterance) it seems to do so in virtue of the fact that the earlier utterance of the proper name 'Albert' referred to Albert. Thus, there appears to be an anaphoric link between the utterances 'Albert' and 'the cat', which means that 'the cat', in this example, is an anaphoric definite description utterance.[24]

Anaphoric definite description utterances provide a data point that semantic analyses of definite descriptions must accommodate. Any such analysis must explain how anaphoric definite description utterances come to refer to their referents in virtue of their antecedents referring to those referents. Furthermore, the principle of parsimony suggests that an ideal semantic analysis of definite descriptions would account for anaphoric definite description utterances without positing any additional principles.

The salience analysis can accommodate these cases in a parsimonious fashion. The idea is to "subsume the anaphoric use under the ... salience use of definite [descriptions]" (von Heusinger, 2004, p. 315). To do so, we first have to recall the principle I noted in Section 1: mentioning or pointing to an entity raises its salience. Under the salience analysis, an utterance of 'the cat' will refer to the most salient cat in the context. We can suppose that, in **(2)**, no cats were salient before David spoke (given that no-one has

[23] Note that I made three changes to Lewis' original example. Firstly, I eliminated any other utterances or salience-raising activities so that I could focus on a single sentence. Secondly, I changed 'our New Zealand cat' to 'Albert', as phrases of the form 'our *F*' are often themselves interpreted as definite descriptions (i.e. 'the *F* of ours') and it would confuse the issue to have a definite description as both antecedent and anaphor. Third, I changed the two sentences of Lewis' example into one sentence, as intra-sentential anaphoric links are a better example of the phenomenon I am interested in that inter-sentential ones. I'll use 'David' to refer to the speaker in **(2)** so he is not confused with Lewis himself.

[24] Incidentally, David's utterance of 'he' is also anaphorically linked to his utterance of 'Albert'.

mentioned any cats previously in the context, and no cats are otherwise relevant). Thus, before the sentence was spoken, all the cats had equal salience, meaning that there was no maximally salient cat, so any utterance of 'the cat' at this point in time would have had no referent. According to the principle noted above, by uttering 'Albert' David raised Albert's salience in the context. Thus, it would no longer be the case that all cats had equal salience in the context. Instead, Albert will be the maximally salient cat in the context after David's utterance of 'Albert'. This means that David's later utterance of 'the cat' will refer to Albert, as Albert is the maximally salient cat in the context at that point in time. Thus, the salience analysis both correctly predicts the referent of anaphoric definite description utterances, and explains how such utterances get their referents in virtue their antecedents referring to the referents (namely, the anaphora refer to their referents in virtue of their antecedents making those referents more salient by mentioning to them).

On the face of it, it may not seem like this is a parsimonious approach to these cases, as it seemed to require the postulation of a new principle. I will suggest that this is a parsimonious response, because this principle is not a new principle, but rather (like the salience principles I noted earlier) it is justified by our intuitive notion of salience. It seems intuitively clear that mentioning something, or pointing to it, will make it more salient to those who hear it mentioned or see it pointed to.

This principle, as I have presented it so far, is quite vague, in so far as I have not specified to *what degree* the salience of an entity is raised by mentioning or pointing to it. I will suggest that mentioning or pointing to an entity raises it to the *highest* level of salience, so that every other entity is less salient than it is. We can see why this stipulation is important by considering Lewis' (1979, p. 348, emphasis mine) original example:

> In the room is a cat, Bruce, who has been making himself very salient by dashing madly about. He is the only cat in the room, or in sight, or in earshot. I start to speak to you: 'The cat is in the carton. The cat will never meet *our other cat*, because our other cat lives in New Zealand. Our New Zealand cat lives with the Cresswells. And there he'll stay, because Miriam would be sad if *the cat* went away.'

Before Lewis' utters 'our other cat' for the first time, Bruce is the uniquely maximally salient cat. Suppose Lewis' utterance of 'our other cat' raised Albert's salience, but only slightly. If that were the case, then Lewis' final utterance of 'the cat' would not necessarily refer to Albert according to the salience analysis, and we would then have no explanation of why it does refer to Albert. In contrast, Lewis' utterance of 'our other cat' raising Albert's salience to the highest level would *guarantee* that Lewis' final utterance of 'the cat' would refer to Albert. This is why the stronger version of the principle is needed.

2.2 Static and Dynamic Semantics

On a standard view of semantics, the reference of a term utterance is determined by the meaning of the phrase and (in at least some cases) the context of the utterance. Within formal theories of meaning—following that of Kaplan (1980)—this idea is embodied in an understanding of the meanings of phrases as functions from the contexts of use of the phrase to the denotation of the phrase. Note that, on this understanding, the relationship between reference and context is one-way, in the sense that the context affects reference but referring does not affect the context. This understanding is inconsistent with the principle I noted above, as that principle holds that referring to something can affect the salience of the context, and therefore referring does affect the context.

There is no reason why we have to stick with this understanding of the relationship between reference and context. Kaplan seems to have adopted this approach only because he was not concerned with the effects that referring could have on contexts. However, any attempt to alter a Kaplan-style semantics to take into account the effects that referring can have on contexts will quickly run into problems. The biggest problem is the following: Kaplan-style semantics are entirely *vertical*, in the sense that we start by trying to determine the truth-value of a sentence relative to a particular context, and in the process of doing so we find that we have to determine the denotation of the subsentential phrases of that sentence relative to *that same context*, and use those denotations to determine the truth-value of the sentence. We start at the top with the sentence, work our way down to the smallest phrases, and then work our way back up to the sentence. The up-shot of this is that, on standard Kaplan-style semantics, we must evaluate every phrase within a sentence relative to the *same* context. This rules out the possibility that referring could affect the context, as referring is something that occurs *during* the utterance of a sentence. The explanation of **(2)** given above would not be possible, as it requires the utterances of 'Albert' and 'the cat' to be evaluated relative to different contexts, despite being parts of the same sentence.

These vertical, Kaplan-style semantics can be called *static*, in so far as they do not allow for *context change* during the evaluation of a sentence utterance. No static semantics will allow for the explanation of anaphoric definite description utterances I gave above. Luckily, *dynamic* theories of meaning exist, which will allow for context change. In the next subsection, I will present such a theory, and in the following subsection I will use it to formalize this explanation of anaphoric definite description utterances.

2.3 Dynamic Semantics with Choice Functions

Stalnaker (1978, p. 78) noted that "acts of assertion affect, and are intended to affect, the context, in particular the attitudes of the participants in the situation." With this idea in mind, "dynamic semantic theories like Discourse Representation Theory (DRT) ... or Dynamic Predicate Logic (DPL) ... assume that the meaning of a sentence is identified with its context change potential" (von Heusinger, 2004, p. 318, my emphasis). To understand the meaning of a sentence, on these views, one must understand how an utterance of that sentence would affect a given context. Already, it should be clear how such dynamic semantics will allow for the context change needed for the salience explanation of anaphoric definite description utterances, where static semantics do not.

However, it is not yet clear how one would represent the kinds of context change invoked in the salience explanation within a dynamic semantics. Furthermore, it also unclear how context change potentials could do the other work we require from meanings, such as providing truth conditions for sentences. To see how these things are both possible, I will have to begin introducing some of the technical machinery of von Heusinger's *Dynamic Semantics with Choice Functions* (DSCF).

In most theories of dynamic semantics, contexts "contain two kinds of information: information about the world, and discourse information. The information about the world is relevant for the truth conditions, while the information about the discourse restricts anaphoric relations" (von Heusinger, 2004, p. 318). To represent information about the world, dynamic semantics borrow an idea from Stalnaker. Stalnaker (1978, p. 84-85) defines a *context set* as "set of possible worlds recognized by the speaker to be the "live options" relevant to the conversation." The possible worlds in a context set are those that the participants of a conversation presuppose *could be* the actual world, and as such the context set contains information about how the world is as far as the participants of the conversation are concerned.[25]

Under the salience analysis, the discourse information which restricts anaphoric links is information about the salience of a context. As such, a dynamic formalization of the salience analysis would require a contextual representation of salience. DSCF is such a formalization, and it utilizes "one global choice function which stands for the salience structure of the context" (von Heusinger, 2004, p. 316). *Choice functions* are functions that take a set as their argument and return a member of that set as their value. In other words, a function Φ is a choice function if and only if, for every set X in the functional domain of Φ,

[25]This understanding of contexts differs from the Kaplan-style understanding I have been using so far, which contains a single possible world parameter. I will explain this discrepancy below.

$\Phi(X) \in X$.[26] The entity $\Phi(X)$ can be called the *choice element* of X relative to Φ. In DSCF, "the salience structure is semantically reconstructed by a global choice function that assigns to each set one (most salient) element" (von Heusinger, 2004, p. 309). In other words, the salience choice function Φ_c of a context c represents the salience structure of c, in so far as $\Phi_c(X)$ is understood to be the uniquely most salient member of X, for all subsets X of the domain of c. The context c will then be the pair $\langle w_c, \Phi_c \rangle$, reflecting the fact that the choice function is "supplied from a global parameter of discourse" (von Heusinger, 2004, p. 311). Recalling that $\|P\|^c$ is the denotation of a phrase P in a context c, DSCF formalizes the salience analysis as follows:[27]

$$\text{(SC)} \qquad \|\text{the } F\|^c = \Phi_c(\|F\|^c)$$

One problem with this approach should already be noticeable, namely that not every subset of the domain has a most salient member. In particular, empty sets do not have maximally salient members, as they have no members. Traditionally, choice functions do not include the empty set in their f-domains, as the condition $\Phi(X) \in X$ cannot hold when X is empty. However, this just creates another problem, as (SC) requires Φ_c to have *every* possible subset of the context's domain as part of its f-domain if it is to apply in every case, including cases of empty definite description utterances (i.e. those where $\|F\|^c = \emptyset$).

Both of these problems, von Heusinger (2004, p. 327en) recognizes, can be resolved if we "assume that the choice function which represents the global salience parameter is a partial function." A function is *partial* when it only has a defined value for *some* of the members of its f-domain. A *partial choice function* is therefore a function where $\Phi(X) \in X$ holds for all those sets on which $\Phi(X)$ is defined. Note that these functions can have the empty set in their f-domain, although they cannot be defined for empty sets, as the condition $\Phi(X) \in X$ cannot hold when X is empty.

A *salience partial choice function* Φ_c of a context c is then a partial choice function which is undefined for all and only the subsets of the domain which do not have a uniquely most salient member, and which returns the most salient member of a subset of the domain for those subsets with most salient members. We can then revise (SC) to take non-referring definite description utterances into account as follows:

[26] I will use the term 'functional domain' or 'f-domain' so that the domains of functions are not confused with the domains of contexts.

[27] This follows von Heusinger (2004, p. 321).

$$\text{(SP)} \qquad \|\text{the } F\|^c = \begin{cases} \Phi_c(\|F\|^c) & \text{if } \Phi_c(\|F\|^c) \text{ is defined} \\ \text{nothing} & \text{otherwise} \end{cases}$$

Is a salience analysis with partial choice functions adequate? It gets the correct result in the case of empty definite description utterances: If an utterance of 'the F' is empty in a context c, then $\|F\|^c = \emptyset$. Φ_c must then be undefined for $\|F\|^c$, so the utterance has no reference, which is the correct result. I will leave further discussion of the adequacy of this approach until the next section. For simplicity, I will refer to salience partial choice functions as *salience functions* for the rest of the paper.

To completely formalize the explanation of anaphoric definite description utterances provided by the salience analysis, it will also be necessary to formalize the notion of salience change it invokes. Von Heusinger (2004, p. 320) formally represents the phenomenon of an entity being 'raised' to maximal salience (as occurs when an entity is referred to) using the *salience function update function*, which is defined below (for a salience function Φ, entity x and set X):

$$upd_1(\Phi, x, X) = \Phi' \text{ such that } \forall X' \begin{cases} \Phi'(X') = x & \text{if } X' = X \text{ and } x \in X' \\ \Phi'(X') = \Phi(X') & \text{otherwise} \end{cases}$$

The idea behind this function is that $upd_1(\Phi, x, X)$ is the result of taking Φ and altering it so that x is the uniquely most salient member of X.

Von Heusinger's method of representing of salience in terms of partial choice functions differs from Lewis' (1973) method of representing of salience in terms of weak orderings. One might ask how important this difference is. In Section 3, I will argue that this difference is important, and Lewis' method should be favored.

2.4 DSCF in Action

As you will recall, the problem with static semantics was that it was entirely vertical, and therefore could not account for context change during the utterance of a sentence. Thus, what the salience approach to anaphoric definite description utterances requires is a *horizontal approach* to salience, where the contextual evaluation of a phrase can be affected by the evaluation of phrases to the left of it in the

sentence. Having laid the groundwork by outlining what von Heusinger takes a context to be, I will now show how a horizontal semantics is possible. I will take a simplified version of von Heusinger's DSCF as an example.[28]

Let's define a *ccp* function so that $ccp(P)$ is the context change potential (i.e. meaning) of the phrase P. A context change potential can be understood as function from contexts to contexts, so that $ccp(P)(c)$ is the context that results from an utterance of P in context c. In DSCF, an utterance can change a context in two ways: either it can change the context set of possible worlds, or it can change the salience choice function (or both). Thus, it will be simpler to present context change potentials if we split them into two parts: a *world change potential*, which tracks changes to the context set, and a *salience change potential*, which tracks changes to the salience function. If we define $wcp(P)$ as the world change potential of a phrase P, and $scp(P)$ as the salience change potential of P, then $ccp(P)(c) = \langle wcp(P)(c), scp(P)(c)\rangle$.

To understand how world change potentials function, let's go back to Stalnaker again. Stalnaker (1978, p. 86) says that "to make an assertion is to reduce the context set" so "that all of the possible situations incompatible with what is said are eliminated." On this view, the world change potential of an assertion can be seen as a filter, which allows through possible worlds which are compatible with the assertion, and blocks those which aren't. Thus, "the essential effect of an assertion is to change the presuppositions of the participants in the conversation by adding the content of what is asserted to what is presupposed" (Stalnaker, 1978, p. 86). Note that this world change potential, *qua* meaning, is not spelled out directly in terms of truth conditions. This is because, for Stalnaker (1978, p. 85), the main purpose of assertion is not to say true things, but "distinguish among alternative possible ways that things may be." However, we can use this notion of a world change potential to determine truth conditions for a sentence, in so far as a sentence φ is true in a possible world w if and only if an utterance of φ in a context containing w in its context set would not eliminate w from the context set. In line with this, I will define world change potentials not in terms of contexts with context sets, but rather in terms of contexts c with single worlds w_c, so that φ is true in c if and only if $wcp(\varphi)(c) = w_c$, and false in c if and only if $wcp(\varphi)(c)$ has no value.[29] The extension of this notion to context sets should be clear.

[28] Von Heusinger's DSCF has additional complications due to his aim of also representing the salience change potential of indefinite descriptions. These additions are unnecessary for my purposes.

[29] This is how single-world Kaplan-style contexts can be reconciled with Stalnaker-style context sets. Note that what is defined here are not strictly speaking truth conditions, in so far as they define when an assertion of a sentence φ is true in a context, rather than when the sentence φ is true at a world. Thus, on this model of truth conditions, we cannot answer the question of whether φ would be true in a world with no language-users, as there do not exist contexts without language-users. I do, however, think that the notion of truth conditions defined here is close enough to the original notion that we can understand how context change potentials could function as meanings.

In contrast with world change potentials, salience change potentials do not bear directly on truth conditions (although they do have an indirect effect, as I will show later). The main thing to note is that, in dynamic semantics like DSCF, subsentential phrases also have salience change potentials. We can see in the abstract how two salience change potentials (associated with constituent phrases) can be combined into a complex salience change potential (associated with a more complex phrase or a sentence) by considering how two changes, one after the other, form one big change. Centrally, later phrases are evaluated relative to the context *after earlier phrases have changed it*, not the context that existed before the sentence was uttered. Thus, we can see how context change is possible on this account.

To demonstrate how a system of context change potentials can formalize both a horizontal semantics and the salience approach to anaphoric definite description utterances, consider the following language fragment:

- Proper Names: 'Albert'
- Determiners: 'the'
- Predicates: 'cat', 'is in NZ', 'will stay there'
- Connectives: 'and'

I will define $\|\text{Albert}\|^c$, for all contexts c, as **Albert** (i.e. the very cat Albert). The syntax of this fragment is that of standard English. The meaning of the phrases then formed can be spelled out in terms of their world and salience change potential. We can start with the only atomic sentences that can be formed in the fragment, which are sentences of the form 'T F', where 'T' is a term (either 'Albert' or 'the F') and 'F' is a predicate.[30]

$$wcp(T\ F)(c) = \begin{cases} w_c & \text{if } \|T\|^c \in \|F\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(T\ F)(c) = scp(T)(c)$$

[30]Treating definite descriptions as terms in dynamic semantics is an innovation of von Heusinger (2004, p. 321).

The world change potential of '$T\ F$' holds that a possible world is compatible with an utterance of '$T\ F$' in that world if and only if the referent of 'T' in that world is a member of the extension of 'F' in that world.[31] These are the standard truth conditions for a subject-predicate sentence. Note that, where 'T' is a definite description, $\|T\|^c$ is defined in accordance with (SP). Thus, we can see this truth condition as implementing the salience analysis. Also note that the world change potential of such a sentence is not defined in terms of the world change potentials of its components. Thus, there will be no need to define world change potentials for subsentential phrases.

The salience change potential of '$T\ F$' is identical to that of 'T'. I am making a simplifying assumption here that utterances of predicates do not themselves affect salience. This means that no salience change potential will need to be defined for predicates.[32]

Now, let's consider the salience change potential of terms, starting with 'Albert':

$$scp(\text{Albert})(c) = upd_1(\Phi_c, \mathbf{Albert}, \|\text{cat}\|^c)$$

This definition says that an utterance of 'Albert' will update the salience choice function so that Albert is the most salient cat. This is in accordance with the principle suggested above, that mentioning an entity makes that entity maximally salient. One might wonder why the salience is updated so that Albert becomes the most salient *cat*, and not the most salient animal, or New Zealander, or New Zealander cat, or any of the other categories Albert is a member of. In the next section, I will suggest that Albert is in fact the most salient member of all these categories after an utterance of his name, and that it is a weakness of this approach that this fact is not recognized. However, I will keep using this notion of salience updating as an approximation, to display how DSCF works, before presenting a more accurate notion later.

We can now see how DSCF works for the example sentence, 'Albert is in NZ':

$$wcp(\text{Albert is in NZ})(c) = \begin{cases} w_c & \text{if } \|\text{Albert}\|^c \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

[31] I will take the condition $\|T\|^c \in \|F\|^c$ to be unfulfilled if $\|T\|^c$ is undefined (i.e. if T has no referent). If one instead thought that assertions containing non-referring terms were neither true nor false, one would need a different semantical apparatus, as the only distinction this apparatus provides is either eliminating or not eliminating a world. In interests of simplicity I will leave this possibility aside.

[32] Von Heusinger (2004, p. 322) also only considers the salience change potential of terms and not predicates.

$$wcp(\text{Albert is in NZ})(c) = \begin{cases} w_c & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ})(c) = scp(\text{Albert})(c)$$

$$scp(\text{Albert is in NZ})(c) = upd_1(\Phi_c, \mathbf{Albert}, \|\text{cat}\|^c)$$

An utterance of 'Albert is in NZ' will then have two effects on the context. Firstly, it will eliminate any world where the referent of 'Albert', i.e. the cat Albert, is not in the extension of 'is in NZ', i.e. the set of entities in NZ. In other words, this utterance is true if and only if Albert is in NZ, which matches the intuitive truth conditions of the sentence. Secondly, the utterance of 'Albert is in NZ', by virtue of containing an utterance of 'Albert', makes Albert the most salient cat in the context, which accords with the claims made about salience above.

Next, we can consider the salience change potential of a definite description 'the F':

$$scp(\text{the } F)(c) = \begin{cases} upd_1(\Phi_c, \|\text{the } F\|^c, \|F\|^c) & \text{if } \|\text{the } F\|^c \text{ is defined} \\ \Phi_c & \text{otherwise} \end{cases}$$

This definition suggests that how a definite description utterance changes contextual salience depends on whether it refers. In the non-referring case, this definition states that the utterance makes no change to salience. This is in accordance with the principle stated above that mentioning an entity raises its salience. A non-referring definite description utterance mentions nothing, and therefore makes nothing more salient.

In the referring case, this definition states that an utterance of 'the F' will update the salience choice function so that the referent of 'the F' will be the most salient F. Of course, as noted by von Heusinger (2004, p. 317), under the salience analysis, the referent of 'the F' is *already* the most salient F. Thus, on according to this definition of the salience change potential of definite descriptions, referring utterances of definite descriptions *also* do not affect salience at all. In other words, $scp(\text{the } F)(c) = \Phi_c$. This is not strictly incompatible with the claim made above that mentioning something makes it more salient, as we could hold that an utterance of a definite description would increase its referent's salience if its referent

wasn't already maximally salient. We could then say that utterances of definite descriptions do update the salience ranking, albeit *trivially* so. The claim that an utterance of, for example, 'the cat' only makes its referent the maximally salient cat seems more justified than the claim that an utterance of 'Albert' only makes its referent the maximally salient cat, but I think that it is still incorrect, as I will discuss in the next section.

Now let's consider the example sentence 'The cat is in NZ':

$$wcp(\text{The cat is in NZ})(c) = \begin{cases} w_c & \text{if } \|\text{the cat}\|^c \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$wcp(\text{The cat is in NZ})(c) = \begin{cases} w_c & \text{if } \Phi_c(\|\text{cat}\|^c) \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{The cat is in NZ})(c) = scp(\text{the cat})(c)$$

$$scp(\text{The cat is in NZ})(c) = \Phi_c$$

An utterance of 'The cat is in NZ' then has the following effects on the context: Firstly, it will eliminate any world where either 'the cat' does not refer, or the referent of 'the cat', i.e. the most salient cat, is not in the extension of 'is in NZ'. In other words, this utterance is true if and only if there is a most salient cat and it is in NZ, which matches the truth conditions of the sentence according to the salience analysis. Note that this is a case where salience affects truth conditions. Secondly, the utterance of 'The cat is in NZ', by virtue of containing an utterance of 'the cat', trivially makes the most salient cat in the context the most salient cat in the context (i.e. leaves the salience choice function as it is).

Finally, we can look at sentences of the form 'φ and ψ', where φ and ψ are sentences:

$$wcp(\varphi \text{ and } \psi)(c) = wcp(\psi)(\langle wcp(\varphi)(c), scp(\varphi)(c)\rangle)$$

$$scp(\varphi \text{ and } \psi)(c) = scp(\psi)(\langle wcp(\varphi)(c), scp(\varphi)(c)\rangle)$$

The basic idea is that the context that results from uttering 'φ and ψ' is identical to the context that would result from uttering φ and then uttering ψ.[33] To see this in action, consider the sentence 'Albert is in NZ and the cat will stay there'. This is a simplified version of the sentence in **(2)**, and like that sentence there is an anaphoric link between the utterances of 'Albert' and 'the cat'.[34] Thus, using this example, we can see how DSCF can account for the context change necessary for the salience explanation of anaphoric definite description utterance. For convenience, I will abbreviate $upd_1(\Phi_c, \mathbf{Albert}, \|\text{cat}\|^c)$, i.e. Φ_c except Albert is the most salient cat, as $\Phi_c^{\mathbf{A}}$, and the context $\left\langle w_c, \Phi_c^{\mathbf{A}}\right\rangle$, i.e. c except Albert is the most salient cat, as $c^{\mathbf{A}}$.

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$wcp(\text{the cat will stay there})(\langle wcp(\text{Albert is in NZ})(c), scp(\text{Albert is in NZ})(c)\rangle)$$

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} wcp(\text{the cat will stay there})(c^{\mathbf{A}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} w_c & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \text{ and } \mathbf{Albert} \in \|\text{will stay there}\|^{c^{\mathbf{A}}} \\ \text{nothing} & \text{otherwise} \end{cases}$$

[33] Von Heusinger (2004, p. 322) takes a similar approach.

[34] A reading of this sentence with an anaphoric link is somewhat strained. A more natural way of expressing this point would be to say 'Albert is in NZ and *he* will stay there'. I use it as an example because it is much shorter than the example in **(2)**. I think that we can see my suggested reading if, for example, we suppose that we were in a context with Samson the dog, and someone says 'Samson is in Australia, and he will move, while Albert is in NZ, and the cat will stay there'.

$$scp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$scp(\text{the cat will stay there})(\langle wcp(\text{Albert is in NZ})(c), scp(\text{Albert is in NZ})(c)\rangle)$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} scp(\text{the cat will stay there})(c^{\mathbf{A}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) = \begin{cases} scp(\text{the cat})(c^{\mathbf{A}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) = \begin{cases} \Phi_c^{\mathbf{A}} & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

We can then see that an utterance of 'Albert is in NZ and the cat will stay there' will eliminate a possible world from the context set if and only if it is not the case that Albert is in NZ and will stay there. These are the intuitive truth conditions of the sentence *given 'the cat' is an anaphoric definite description utterance*. Thus, DSCF correctly models the anaphoric link in this utterance. Furthermore, we can see DSCF operating here as a horizontal semantics, as the constituent sentence 'the cat will stay there' is evaluated in the context that results from updating the salience choice function on the basis of the utterance of 'Albert' (if it is evaluated at all). Thus, DSCF provides the possibility that the contextual evaluation of a phrase can be affected by the evaluation of phrases to the left of it in the sentence. Finally, the result of uttering this sentence, supposing it is true, is to update the salience choice function so that Albert is the most salient cat, in accordance with the salience principles discussed above.

Thus, we can see how DSCF models the context change necessary for the salience explanation of anaphoric definite description utterances. One can note that this modeling thoroughly utilizes salience functions. This raises the question of whether it is necessary to use the choice-functional method of

salience representation to model this kind of context change. In the next section, I will show that it is *not* necessary, by creating a version of DSCF which uses salience orders rather than salience functions, and which still models the needed context change.

3 The Choice-Functional and the Order-Theoretic Methods of Salience Representation

3.1 Comparing the Methods

One key difference between the salience analyses offered by von Heusinger (2004) and Lewis (1973) is in how they represent salience. Lewis represents salience using weak orderings while von Heusinger uses partial choice functions. On the face of it, this difference may not seem very important. Weak orderings and partial choice functions both serve to represent the most salient members of subsets of the domain (if there is one). We might suppose that the choice between them is merely a technical one, like the choice between the order-theoretic and system of spheres methods we looked at in Section 1. In this section, I will show that this is not the case. The order-theoretic method can explain certain features of contextual salience which the choice-functional method cannot, and for that reason the order-theoretic method should be favored.[35] One might then think that, even if the choice-functional method is a worse method of representing salience, we have to trade that off with the advantages we get from being able to model the salience change necessary to explain anaphoric definite description utterances. To show that this is not the case, I will introduce Dynamic Semantics with Orderings, an alteration of DSCF using salience orders, to show that no trade-off is necessary, as we can model salience change as easily with salience orders as with salience functions.

Before I can make the argument that the order-theoretic method should be favored, I must first show that the two methods are not, in fact, equivalent, in the sense I discussed in Section 1. If these two methods are equivalent then, like the order-theoretic and system of spheres methods, any choice between the two must be merely technical. Recall that two methods of salience representation are equivalent if and only if there is a one-to-one relation between representations in one method and equivalent representations in the other. Thus, to show that the choice-functional and the order-theoretic methods are not equivalent, I need to show that it is not the case that, for every salience order $\geq_i$ we can define an equivalent salience function Φ_i, and vice versa.

[35] Note that I am not claiming that the the choice-functional method is inadequate (or that it could not be made adequate with the stipulation of further principles). Rather, I am claiming that *even if* the choice-function were adequate it still could not explain facts about salience that the order-theoretic method can.

It is true that, for each salience order $\geq_i$ we can define an equivalent salience function Φ_i. Let's call the set of entities $\geq_i$ is defined over U_i. For every subset X of U_i, $\Phi_i(X)$ will be defined if and only if $\max(\geq_i, X)$ is a singleton. Supposing $\Phi_i(X)$ is defined, its value will be the single member of $\max(\geq_i, X)$. The equivalence of $\geq_i$ and Φ_i should be easy to see by comparing (SO) and (SP).

It is not, however, possible to define an equivalent salience order for every salience function. We can see this by considering two classes of counterexamples. Firstly, suppose we have a salience function Φ_a and an entity x such that $\Phi_a(\{x\})$ is undefined. This would mean that, in a context with the salience function Φ_a and a predicate 'F' with the extension $\{x\}$, an utterance of 'the F' would have no referent. However, *every* salience ordering is such that 'the F' would refer to x in such a context. This is because, for every salience ordering $\geq$ and singleton $\{x\}$, $\max(\geq, \{x\}) = \{x\}$, as I showed in Section 1. Secondly, suppose we have a salience function Φ_b and three entities x, y and z such that $\Phi_b(\{x, y, z\}) = x$ but $\Phi_b(\{x, y\}) = y$. This would mean that, in a context with the salience function Φ_b, a predicate 'G' with the extension $\{x, y, z\}$ and a predicate 'H' with the extension $\{x, y\}$, an utterance of 'the G' would refer to x but an utterance of 'the H' would refer to y. However, supposing we define a salience ordering $\geq_b$ for this context, for 'the G' to refer to x it is necessary that $x >_b y$, yet for 'the H' to refer to y it is necessary that $y >_b x$. Thus, such a $\geq_b$ is impossible. The issue, essentially, is that, a) salience functions need not be defined for singletons, which means that (unlike for salience orderings) proper definite description utterances don't have to have referents, and b) the choice element for a set can be defined independently of the choice elements of any of its subsets, while which element is maximal in a set can affect which elements are maximal in its subsets.

Thus, the order-theoretic and choice-functional methods are not equivalent. We are then left with the question of which method is more theoretically virtuous. On the face of it, we might suppose that the choice-functional method is better, as there are (seemingly) possible salience rankings which it can represent that the order-theoretic method cannot. I think this impression is *wrong*. I will argue that the greater generality of the choice-functional method is, in fact, a *weakness*, as it means that the choice-functional method cannot *explain* certain features of salience that the order-theoretic method can. In the next three subsections of this section, I will present my argument that the order-theoretic method better represents salience.

3.2 The Singleton Condition

As I noted above, nothing I have said so far excludes salience functions from being undefined for singleton arguments. This is a major problem, as it means that the salience analysis using this method does not entail (UP): Proper definite description utterances refer to the single member of their restrictor sets. As such, this method does not provide for an adequate analysis of definite description reference. This problem can be fixed by introducing the following *singleton condition*: all salience choice functions are defined for singletons. The singleton condition, together with the definition of partial choice functions which holds that $\Phi(X) \in X$ for every X where $\Phi(X)$ is defined, entail that for every singleton $\{x\}$, $\Phi(\{x\}) = x$.

However, the singleton condition raises a further problem, which is that it seems *unjustified*. There does not seem to be a more general intuitive principle of salience which explains why the singleton condition holds. This is problematic for a salience analysis utilizing the choice-functional approach for two reasons. Firstly, the singleton condition's lack of justification makes it *ad hoc*, which means that it cannot tell us *why* all proper definite description utterances refer to the single member of their restrictor sets. This takes away some of the *explanatory strength* of the view. Secondly, given the singleton condition is not otherwise justified, it then adds to the *theoretical commitments* of the analysis, making it a more *complicated*, and therefore worse theory.

In contrast, the salience analysis using the order-theoretic method does not have this problem. As we saw in Section 1, given all salience orderings are reflexive, it must be the case that proper definite description utterances refer to the single member of their restrictor sets. Thus the reflexivity of salience orderings *explains* the truth of (UP) and the totality of salience orderings *explains* their reflexivity. The salience analysis using the order-theoretic method is therefore more explanatorily powerful and theoretically parsimonious than the one using the choice-functional method.

One final note on this topic. If we held that all salience functions were *total*, then proper definite description utterances would not raise the same problem for the choice-functional method. This is because it would then be the case that all salience functions are defined for all singletons, and (as I noted above) if salience functions are defined for singletons then (UP) is true. There would, in that case, be an explanation of why (UP) is true (namely, because all salience functions were total). However, as we saw in Section 2, this principle cannot be justified as it is not true. It would have the consequence that all definite description utterances refer, which is clearly false. Thus, this 'solution' to the problem of singletons is not a solution at all.

3.3 The Monotonic Decreasing Condition

Consider this example, adapted from Lewis (1973, p. 114): **(Example 3)** 'David and Robert are walking past a piggery. One pig, Percy, is conspicuously running around, while the other pigs are standing still further away. David says 'The pig is grunting'.' If we ask the question, what pig was David referring to, the answer seems to be Percy. This much is predicted by the salience analysis. Now, let's alter **(3)** so that David instead says 'The pig with floppy ears is grunting'. Does David still refer to Percy in this case? If Percy does not have floppy ears then the answer is simple: an utterance of 'the pig with floppy ears' can only refer to a pig with floppy ears, so the answer must be no. However, if Percy *does* have floppy ears, the answer also seems to be simple: yes. If David's utterance in the original example referred to Percy, that could only be because something about Percy made him more salient than every other pig. These same qualities must make him more salient than every other floppy-eared pig, as all floppy-eared pigs are in the set of all pigs. More generally, we can say that definite descriptions are *monotone decreasing* on the salience analysis: if 'the F' refers to x in c, and $\|G\|^c \subseteq \|F\|^c$, and $x \in \|G\|^c$, then 'the G' also refers to x in c.

This property of salience is not predicted by the choice-theoretic method as it has been presented so far. This is because the choice element for a set can be defined independently of the choice elements of any of its subsets. For example, suppose we defined a salience function Φ_c so that $\Phi_c(\|\text{pig}\|^c) = \textbf{Percy}$ but $\Phi_c(\|\text{floppy-eared pig}\|^c) \neq \textbf{Percy}$ (where **Percy** is that very pig Percy). This is completely consistent, given what has so far been said, with Percy being a member of $\|\text{floppy-eared pig}\|^c$. To make sure that the choice-theoretic method makes the correct predictions in cases like **(3)**, we need to introduce the *monotonic decreasing condition* on salience functions: if Φ is defined for X and Y and $Y \subseteq X$ and $\Phi(X) \in Y$ then $\Phi(Y) = \Phi(X)$.

Like the singleton condition, the monotonic decreasing condition also seems to be unjustified. Again, there does not seem to be a more general intuitive principle of salience which explains why it holds. As such, the principle is *ad hoc*, which means that it cannot tell us *why* definite descriptions are monotone decreasing, thus again reducing the explanatory strength of the view. Furthermore, the principle also increases the theoretical commitments of the analysis.

The order-theoretic method does not have these problems, as it predicts that definite descriptions are monotone decreasing. To see this, consider that if 'the F' refers to x in c, then $x >_c y$ for all y other than x in $\|F\|^c$. If $\|G\|^c \subseteq \|F\|^c$, and $x \in \|G\|^c$, then it must also be the case that $x >_c y$ for all y other than x in $\|G\|^c$. This means that 'the G' also refers to x in c. We can also see how the order-theoretic method

explains why definite descriptions are monotone decreasing, as it is a general principle of order theory that the maximal element of a set X is also the maximal element of any subset of X which contains it. Note that this explanation does not appeal to any of the special conditions placed on salience ordering, as it would apply not matter what kind of orderings salience orderings were.

3.4 Updating Salience

In Section 2, I raised the issue that it was unclear why a referring utterance of 'the F' only made its referent the most salient F, rather than the most salient member of any of the other categories it is a member of. Furthermore, it seemed unclear which category, or categories, the referent of an utterance of a proper name should be raised to be most salient member of. In this subsection, I will suggest these confusions arise from the wrong way of thinking about salience change, and that in fact mentioning an entity makes it the most salient member of *all* of the categories it is a member of. Furthermore, I will suggest that, unlike the choice-functional method, the order-theoretic method provides an explanation of this fact, once again showing it to be the better method of salience representation.

To see why this is the wrong way to think about salience change, consider the following example from Hemingway's "A Clean, Well-lighted Place" quoted by von Heusinger (2004, p. 325, his emphasis):

> **(Example 4)** "**The waiter who was in hurry came over.** "Finished," **he** said ...
> "Another", said the old man. "No, finished." ***The waiter*** wiped the edge of the table with a towel and shook **his** head."

Suppose these sentences were uttered in a context H where there was a most salient waiter who was in a hurry, who we can call Larry.[36] Intuitively, in **(4)** the utterance of 'the waiter' refers to the same entity as the utterance of 'the waiter who was in a hurry', namely the waiter in a hurry. Thus, the salience analysis should predict an anaphoric link between these two utterances.[37] However, such a link is not predicted by DSCF. According to that system, an utterance of 'the waiter who was in a hurry' in H will change the salience function of H to $upd_1(\Phi_H, \|\text{the waiter who was in a hurry}\|^H, \|\text{hurrying waiter}\|^H)$". This function is identical to Φ_H except that Larry will be the most salient hurrying waiter. But Larry is *already* the most salient hurrying waiter, so the salience function of the context after the utterance of 'the waiter who was in a hurry' is just Φ_H. This is an instance of the fact noted in Section 2 that utterances of definite

[36] I will ignore any issues of tense here.

[37] Incidentally, there is also an anaphoric link between the utterances of 'the waiter who was in a hurry' and 'he', and between 'the waiter' and 'his'.

descriptions on this account only trivially affect salience rankings. However, this fact means that the anaphoric link between the utterances of 'the waiter who was in a hurry' and 'the waiter' cannot be explained. The salience explanation of anaphoric definite description utterances holds that such anaphoric links are established in virtue of the antecedent raising its referent's salience so that the anaphor refers to it. However, the utterance of 'the waiter who was in a hurry' in **(4)** does not raise its referent's salience, and the utterance of 'the waiter' need not refer to that referent, supposing there is some *non-hurrying* waiter (let's call him Moe) who is more salient than Larry.[38]

DSCF must therefore be amended to account for such cases. Von Heusinger (2004, p. 326) suggests (in a related case) that "it is obvious that ... *the younger waiter* changes the context in a way that its referent is not only the most salient younger waiter (trivially), but also that its referent is the most salient waiter (at all)." On the basis of this, von Heusinger (2004, p. 326, my emphasis) holds that "definite descriptions ... change the context by raising new referents to the most salient ones for the set they describe *as well as some supersets*." Thus, we could hold that the utterance of 'the waiter who was in a hurry' does not just (trivially) make Larry the most salient hurrying waiter, but also the most salient waiter in general. This would explain the anaphoric link in **(4)**.

This raises the question of *which* supersets of its restrictor set does the referent of a definite description utterance become the most salient member of, after the utterance. Von Heusinger does not answer this question. I will argue that the referent of a definite description utterance becomes the most salient member of *every* superset of the description's restrictor set, after the utterance. To see why this is the case, consider the following example: 'Max pushed the book and the thing fell down'. In this example, the utterances of 'the book' and 'the thing' seem to have an anaphoric link. To explain this case, we will have to hold that the utterance of 'the book' makes its referent the most salient thing. The set of things seems to contain, well, everything. Given the monotonic decreasing property, the referent of the utterance of 'the book' must be the most salient member of every subset of the set of things it is a member of, which includes every superset of the set of books. To explain cases like this, we must hold that uttering a definite description makes its referent the most salient member of every superset of the utterance's restrictor set.

To formalize this conclusion in DSCF, we must replace the salience function update function upd_1 with the following function:[39]

[38]The introduction of the monotonic decreasing condition means that we could explain an anaphoric link between a definite description utterance and a *more* specific definite description utterance, such as 'the waiter who was in a hurry' and 'the *white-haired* waiter who was in a hurry.' This does not help us here, however, as it is completely consistent with the monotonic decreasing condition that Larry be the most salient hurrying waiter without being the most salient waiter (so long as the most salient *non-hurrying* waiter Moe is also the most salient waiter).

[39]This function is based on a function provided by von Heusinger (2004, p. 324), except that von Heusinger seems to incorrectly

$$upd_2(\Phi, x, X) = \Phi' \text{ such that } \forall X' \begin{cases} \Phi'(X') = x & \text{if } X' \supseteq X \text{ and } x \in X' \\ \Phi'(X') = \Phi(X') & \text{otherwise} \end{cases}$$

We can then redefine the salience change potentials for definite descriptions as follows:

$$scp(\text{the } F)(c) = \begin{cases} upd_2(\Phi_c, \|\text{the } F\|^c, \|F\|^c) & \text{if } \|\text{the } F\|^c \text{ is defined} \\ \Phi_c & \text{otherwise} \end{cases}$$

This definition will hold that an utterance of 'the waiter who was in a hurry' in H will make its referent Larry the most salient member of $\|\text{hurrying waiter}\|^H$ *and of every superset of that set*, including $\|\text{waiter}\|^H$. Thus, this reformulated version of DSCF explains the anaphoric link in **(4)**. More generally, note that, on this definition, the salience change potentials of definite descriptions are not trivial, and as such utterances of definite descriptions can alter contextual salience (as in **(4)** for example).

Problems still remain. Firstly, there is the question of the salience change potential of proper names like 'Albert'. After an utterance of 'Albert', which sets should Albert be the most salient member of? Secondly, upd_2 (and, for that matter, upd_1) violates the monotonic decreasing condition. More precisely, the fact that Φ is monotone decreasing is no guarantee that $upd_2(\Phi, x, X)$ is monotone decreasing. To see a counterexample, suppose Φ_H is monotonic decreasing. Then, if $\Phi_H(\|\text{waiter}\|^H) = \textbf{Moe}$ (where **Moe** is that very waiter Moe) and $\Phi_H(\|\text{hurrying waiter}\|^H) = \textbf{Larry}$, it must be the case that Moe is not in a hurry, and that $\Phi_H(\|\text{non-hurrying waiter}\|^H) = \textbf{Moe}$. Let us further suppose that Larry and Moe both have white hair. Then it must be the case that $\Phi_H(\|\text{white-haired waiter}\|^H) = \textbf{Moe}$. Now, let us define $\Phi_{H^{\textbf{L}}}$ as $upd_2(\Phi_H, \textbf{Larry}, \|\text{hurrying waiter}\|^H)$. It will then be the case that $\Phi_{H^{\textbf{L}}}(\|\text{waiter}\|^H) = \textbf{Larry}$, as $\|\text{waiter}\|^H$ is a superset of $\|\text{hurrying waiter}\|^H$. Given $\|\text{white-haired waiter}\|^H$ is a subset of $\|\text{waiter}\|^H$, and Larry is a white-haired waiter, then if $\Phi_{H^{\textbf{L}}}$ was monotonic decreasing then $\Phi_{H^{\textbf{L}}}(\|\text{white-haired waiter}\|^H)$ would have to be Larry. But $\Phi_{H^{\textbf{L}}}$ is identical to Φ_H except for its values relative to $\|\text{hurrying waiter}\|^H$ and its supersets, which does not include $\|\text{white-haired waiter}\|^H$. Thus, $\Phi_{H^{\textbf{L}}}(\|\text{white-haired waiter}\|^H)$ $= \textbf{Moe}$, so $\Phi_{H^{\textbf{L}}}$ is not monotonic decreasing.

We can resolve these two problems by holding that an utterance of a referring definite description makes its referent the most salient member of *every* category it is a member of. We can formalize this by replacing upd_2 with the following function:

define the function so that it makes the referent of the definite description utterance the most salient member of all the *subsets* of the utterances restrictor set, *regardless* of whether the referent is a member of those subsets.

$$upd_3(\Phi, x) = \Phi^{'} \text{ such that } \forall X^{'} \begin{cases} \Phi^{'}(X^{'}) = x & \text{if } x \in X' \\ \Phi^{'}(X^{'}) = \Phi(X^{'}) & \text{otherwise} \end{cases}$$

We can then redefine the salience change potentials for definite descriptions as follows:

$$scp(\text{the } F)(c) = \begin{cases} upd_3(\Phi_c, \|\text{the } F\|^c) & \text{if } \|\text{the } F\|^c \text{ is defined} \\ \Phi_c & \text{otherwise} \end{cases}$$

This redefinition has the advantage that we do not need to select any particular category or categories for the referent of a definite description utterance to be the most salient member of. Furthermore, we can explain the anaphoric link in **(4)**, as an utterance of 'the waiter who was in a hurry' in H will make its referent, Larry, the most salient member of every category Larry is a member of, including the category of waiters. Also, $upd_3(\Phi, x)$ is guaranteed to monotone decreasing if Φ is. To see this, we need to consider two kinds of category F: those where $upd_3(\Phi, x)(F) \neq x$ and those where $upd_3(\Phi, x)(F) = x$. In the first case, $upd_3(\Phi, x)(F) = \Phi(F)$, so this F will not be a counterexample to the monotone decreasing condition. In the second case, if $G \subseteq F$ and $x \in G$ then $upd_3(\Phi, x)(G) = x$, as x will be the most salient member of every category it is a member of, including G, relative to $upd_3(\Phi, x)$, so this F will also not be a counterexample to the monotone decreasing condition. Given there are no possible counterexamples, $upd_3(\Phi, x)$ must be monotone decreasing. Furthermore, note that on this definition of the salience change potentials of definite descriptions it is also the case that utterances of definite descriptions can non-trivially alter contextual salience.

Finally, upd_3 also allows us to redefine the salience change potentials for proper names such as 'Albert' as follows:

$$scp(\text{Albert})(c) = upd_3(\Phi_c, \textbf{Albert})$$

With this definition, we do not need to make a choice about which category or categories Albert will be made the most salient member of after an utterance of 'Albert', as the definition holds that an utterance of 'Albert' will make Albert the most salient member of every category he is member of. This includes the category of cats, so this redefinition of DSCF still explains the anaphoric link in the example sentence 'Albert is in NZ and the cat will stay there'.

Thus, with upd_3, the problems raised by salience updating for DSCF can be solved. Still, we have no explanation of *why* upd_3, and not some other function, is the correct description of how referring utterances raise the salience of their referents. Compare this situation to the order-theoretic approach I discussed in Section 1. There, I claimed that mentioning an entity puts it at the top of the salience ordering. This view has all the advantages discussed so far with regards to upd_3: we do not need to select any particular category or categories for the referent of a definite description or proper name utterance to be the most salient member of. We can explain the anaphoric link in **(4)**, as an utterance of 'the waiter who was in a hurry' will make Larry the most salient thing, and therefore the most salient waiter. The updated salience ordering will be monotone decreasing, as all salience orderings are monotone decreasing, as I showed in the last subsection. Finally, as I will discuss in the next subsection, this view can account for the anaphoric link in the example sentence 'Albert is in NZ and the cat will stay there'.

But this view has a further advantage over the choice-theoretic method, as it can explain *why* mentioning an entity makes it the most salient member of every category it is a member of. If mentioning an entity x makes x more salient than any other entity, then it also makes x more salient than any other entity in a category F x is a member of, and thus makes x the most salient F. As such, the problems raised by updating salience provide another reason why the order-theoretic method has more explanatory power than the choice-functional method.

Ultimately, I think that the three problems for the choice-functional method I have raised here have the same ultimate source. The choice-functional method defines salience as a property entities have *relative to categories*, it so far as it only allows us to answer the question of how salient an entity is by saying whether it is the most salient F, G, etc. This means that any relationship between how salient an entity is in one category and how salient it is in another cannot be explained by the choice-functional method. In contrast, the order-theoretic method defines salience as a property entities have *relative to the whole domain*, which means that it can explain these relationships. Thus, ultimately, is why the order-theoretic method should be favored.

3.5 Dynamic Semantics with Orderings

Hopefully, I have shown how the order-theoretic method is a better method for representing salience than the choice-functional method. However, as we saw in Section 2, explaining anaphoric definite description utterances requires not only representing salience but also modeling salience change. So far, the only way of modeling salience change that I have presented is DSCF, which utilizes salience functions. As

such, despite its weaknesses, we might think that we are forced to use the choice-functional method, lest we be unable to explain anaphoric definite description utterances.

To reject this conclusion, I will show how salience change can be modeled using salience orderings, by altering DSCF so that it utilizes salience orderings. The system I will define I will call *Dynamic Semantics with Orderings* (DSO). The first difference between the two systems that needs to be accounted for is the salience updating function. Rather than the salience function update function upd_3, we need a *salience ordering update function upd*, defined as follows:

$$upd(\geq, x) = \geq' \text{ such that } \forall y, z \begin{cases} y \geq z \text{ and } z \not\geq y & \text{if } x = y \\ y \geq z \text{ iff } y \geq z & \text{otherwise} \end{cases}$$

This function formalizes the claim I made above that mentioning something raises it to the highest level of salience (in every applicable category). Next we need to redefine the world change potentials and salience change potentials of the English fragment I discussed in Section 2. The world change potentials of DSCF do not invoke salience representations directly, and therefore they do not have to be redefined. Note that the world change potentials for subject-predicate sentences are defined in terms of the referent of their subject term. Supposing the subject term is a definite description, this world change potential will utilize the choice-functional principle (SP) in DSCF and the order-theoretic principle (SO) in DSO. However, either way the world change potential *itself* does not need to be redefined, and either way it implements (one version or another of) the salience analysis.

That just leaves the salience change potentials to be redefined. In DSO, salience change potentials must return salience orders, rather than salience functions. As before, I will assume that utterances of predicates do not affect salience, so the salience change potentials of atomic sentences do not need to be redefined, and no salience change potentials need to be defined for predicates. Thus, we can start our redefinition with the salience change potential of 'Albert':

$$scp(\text{Albert})(c) = upd(\geq_c, \mathbf{Albert})$$

This definition entails that an utterance of 'Albert' will make Albert the most salient entity in the context, and therefore the most salient entity in *every* category Albert is a member of. There is then no need to

specify a particular category Albert is the most salient member of. This definition accords with the discussion in the previous subsection. In line with this, we can redefine the salience change potential of 'the F':

$$scp(\text{the } F)(c) = \begin{cases} upd(\geq_c, \|\text{the } F\|^c) & \text{if } \|\text{the } F\|^c \text{ is defined} \\ \geq_c & \text{otherwise} \end{cases}$$

Again, this definition entails that an utterance of 'the F' will make the most salient F the most salient entity in the context, and therefore the most salient entity in every category that entity is a member of. This entity will already be the most salient F, as well as the most salient member of every subset of the Fs (given the monotone decreasing property of salience). However, the entity need not be the most salient member of any *supersets* of the Fs, which means that this salience update need not be trivial. There is no sense here in which the referent of 'the F' is specifically made the most salient F, and as such there is no need to justify this particular choice.

Finally, the salience change potential of conjunctive sentences, like the salience change potential of subject-predicate sentences, makes no direct reference to salience representations, and therefore also does not need to be redefined. We now have the tools to analyze the example sentence 'Albert is in NZ and the cat will stay there'. For convenience, I will abbreviate the context $\langle w_c, upd(\geq_c, \mathbf{Albert})\rangle$ as $c^{\mathbf{O}}$.

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$wcp(\text{the cat will stay there})(\langle wcp(\text{Albert is in NZ})(c), scp(\text{Albert is in NZ})(c)\rangle)$$

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} wcp(\text{the cat will stay there})(c^{\mathbf{O}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$wcp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} w_c & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \text{ and } \mathbf{Albert} \in \|\text{will stay there}\|^{c^{\mathbf{O}}} \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$scp(\text{the cat will stay there})(\langle wcp(\text{Albert is in NZ})(c), scp(\text{Albert is in NZ})(c)\rangle)$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) =$$

$$\begin{cases} scp(\text{the cat will stay there})(c^{\mathbf{O}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) = \begin{cases} scp(\text{the cat})(c^{\mathbf{O}}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

$$scp(\text{Albert is in NZ and the cat will stay there})(c) = \begin{cases} upd(\geq_c, \mathbf{Albert}) & \text{if } \mathbf{Albert} \in \|\text{is in NZ}\|^c \\ \text{nothing} & \text{otherwise} \end{cases}$$

The context change potential defined here is identical to the one defined by DSCF, except that the salience functions are exchanged with salient orderings. Thus we can see that DSO also predicts that an utterance of 'Albert is in NZ and the cat will stay there' will contain an anaphoric definite description utterance of 'the cat'. Furthermore, under DSO, the result of uttering this sentence, supposing it is true, is to update the salience choice function so that Albert is the most salient entity, in accordance with the salience principles discussed above. Thus, we can see that explaining anaphoric definite description utterances does

not require the use of salience functions, and as such we have no reason to favor the choice-functional method over the order-theoretical method.

REFERENCES

Donnellan, Keith. 1966. "Reference and Definite Descriptions." *The Philosophical Review* 75 (3): 281-304.

Frege, Gottlob. 1892. "Über Sinn und Bedeutung." *Zeitschrift fur Philosophie und philosophische Kritik* 100: 25-50. Translated by: Black, Max. 1960. "On Sense and Reference". In *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black, 56-78. Oxford: Blackwell, 1960.

Kaplan, David. 1980. "Demonstratives." In *Themes From Kaplan*, edited by Joseph Almog, John Perry and Howard Wettstein, 481-564. New York: OUP, 1989.

Krahmer, Emiel. 1998. "Chapter 7: Presupposition and Determinedness." In *Presupposition and Anaphora*, Emiel Krahmer, 193-224. Stanford: CSLI, 1998.

Lewis, David. 1973. *Counterfactuals*. Malden: Blackwell, 1973.

Lewis, David. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1): 339-359.

Russell, Bertrand. 1905. "On Denoting". *Mind* 14 (56): 479-493.

Stalnaker, Robert. 1978. "Assertion." In *Context and Content*, Robert Stalnaker, 78–95. Oxford: OUP, 1999.

von Heusinger, Klaus. 2004. "Choice Functions and the Anaphoric Semantics of Definite NPs." *Research on Language and Computation* 2: 309-329.