Informativeness, Category Membership,
and the Distribution of Adjectival Past Participles

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Leah Suzanne van Ameyden van Duym

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Dr. Jeanette Gundel

July 2014

Acknowledgements

First and foremost, I would like to thank my advisor, Jeanette Gundel, for her patience and confidence in me. Her ability to ask just the right question has helped me clarify my thinking and writing on innumerable occasions. I am grateful also to my other committee members, Brian Reese, Hooi Ling Soh, and Michael Kac; their inspiration, guidance, and careful reading have been invaluable.

I am indebted to a large group of wonderful fellow graduate students, who helped this process feel both more doable and more fun. Thanks to Kate, Ellen, Muhammad, Sara, Kaitlin, Eden, Mark, Hiroki, Sharon, Chris, Hiroe, and everyone else who went through this before or with me. Also to Rob, for showing how it's done; to Johanna, for being the first in the family; and to Jen and Daniel, for being themselves.

I am also grateful, of course, to my non-student friends, who helped me keep a foot in the "real world," including Amy, Sam, Heather, Wendy, and Alisa.

Thanks to my family members, who taught me to value scholarly work. I could not and would not have done this without the unwavering support and appreciation of Mom, Ron, Dad, Sherry, Mimi, and Pawpaw. I am grateful for their faith in me, and for all the times they didn't ask when I would be done. My deepest gratitude goes to GraceAnn – my rock, my cheerleader, my other half.

To all those above, and to the many others who have helped me along the way, thank you.

Abstract

Informativeness has been used as an explanatory concept in linguistics and other fields. I discuss previous accounts of informativeness, and propose a model for assessing a component of it, focusing on the distribution of negated and non-negated English adjectival past participles (APPs), like *eaten* and *uneaten*. An utterance is informative to the extent that it increases the cognitive availability of information not fully available via context or category membership, what one knows about an item through knowing its category.

To assess the portion of informativeness related to category membership, or CM-informativeness, I propose a model that uses categories based on qualia (elements of meaning) or linguistic items, which have gradient members whose centrality represents cognitive availability. CM-informativeness relates to centrality: an utterance that associates a category base with some member of that category is more CM-informative if the member is peripheral than if it is central.

The distribution of APPs is explained via the model as follows: an APP occurs more often in non-negated form if its common collocates have categories in which the APP is a peripheral member, because the negated form can then be thought of as a central member. Since an utterance associating a category base with a peripheral member is more informative than one associating it with a central member, and informativeness correlates with frequency, phrases with the non-negated APP occur more often here. Contrastingly, if an APP's common collocates are those with categories in which it is a peripheral member, the reverse occurs.

This model may be generalized to a variety of linguistic forms. For some expressions the CM-informativeness that it assesses is predictive of frequency. This work enriches the conception of informativeness by developing a new characterization and a cognitive model of how one aspect of informativeness may be assessed. This is intended to elaborate on, rather than replace, previous accounts of informativeness.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

   The concepts of information and informativeness have been used in many fields for at least half a century.[1] 'Information' generally refers to the content of a proposition, which can be transmitted and possessed or known by someone; 'informativeness,' on the other hand, is typically used to denote a property of a proposition, message, or part of a message, and involves (but is not necessarily limited to) how much information it contains or supplies. Information and informativeness are often used in the discussion of propositions, or statements that have truth values. However, other elements such as symbols, numbers, phonemes, words, phrases, and functions are also said to carry information and thus have degrees of informativeness (Fogerty & Humes 2010, Wang & Inhoff 2010, van Duym 2010, Shannon 1948). The ideas of information and informativeness are used in a wide variety of disciplines, including semantics, linguistic pragmatics, philosophy, electrical engineering, computer science, and mathematics. Within linguistics, informativeness is most commonly used in accounts, such as Grice's (1975), of the differences between what is said and what is communicated. However, it has also been used in studies of childhood pragmatic development (e.g., Bagassi et al. 2009; Davies & Katsos 2010) and narrative analysis (e.g., Murray 2010), among other topics, and is shown here to be useful in explaining relative frequency of some phrase types.

   Despite their common use, the concepts of information and informativeness often remain ill-defined, and when explications of these two closely related concepts are given, they frequently appear inconsistent both within and across fields. A number of researchers in fields such as speech and hearing sciences, psychology, semantics, and cognitive science have treated informativeness as the degree to which something reduces the number of possibilities for a value, event outcome, etc.[2] For example, given an array of three red shapes, comprising a star, a triangle, and a circle, and the information that John likes exactly one of the shapes, the proposition *John likes the triangle* (which

---

1 E.g., Bruner & O'Dowd (1958) and Shannon (1948).
2 E.g., Fogerty & Humes (2010) on perception, Calhoun (2010) on prosody, Steele & Watkins (2010) on word learning, Syrett et al. (2009) on children's gradable adjective use, Bagassi et al. (2009) on children's use of *some*, and Nieuwland et al. (2010) on processing of scalar statements.

reduces the number of possibilities from three to one) is said to be more informative than *John likes a shape with points* (which leaves two possibilities), and both are more informative than *John likes a red shape*, which is completely uninformative (as it does not eliminate any of the original three possibilities).

This way of measuring informativeness can be thought of as similar to a view, espoused by Horn (1972) and many others, of informativeness as having to do with entailment; put in its simplest form this view states that if A entails B, but B does not entail A, then A is more informative than B. Returning to the propositions above, we can say that *John likes the triangle* entails *John likes a shape with points*, but not vice versa, so again the former statement is viewed as more informative. Other researchers, in fields such as computational linguistics and neurolinguistics, have appealed to the amount (by various measures), correctness, and intuitive relevance of information in assessing informativeness.[3] Many rely to varying extents on a Gricean notion of informativeness; Grice himself did not define 'informative,' and these researchers tend to do the same, but like Grice (1975) they utilize the notion of there being an appropriate level of informativeness in any given conversation.[4] Various other definitions and assessment methods for informativeness abound.

Calhoun (2010), who clearly identifies her use of informativeness as involving both contextual newness and the reduction of possibilities, acknowledges the lack of consistency in uses of 'informativeness.' She notes that "informativeness has been conceptualised in quite different ways in different disciplines" and that "how these different notions relate to each other, or indeed, if they are reducible to a single dimension" is unclear (2010:1101). Such inconsistency is problematic when, as noted by Calhoun (2010), different conceptualizations are not obviously connected or compatible. Compounding this is the fact that many researchers fail to adequately define exactly what they mean by 'informativeness,' leaving the reader to guess which of any number of

---

3 E.g., Stark (2010) on aphasic storytelling, Clarke & Lapata (2010) on document compression, and Murray (2010) on narrative attributes of patients with depression compared with those of patients with Alzheimer's disease.

4 E.g., Davies & Katsos (2010) in applied linguistics, Bagassi et al. (2009) in psychology, and Neiuwland et al. (2010) in cognitive sciences.

possibilities might have been intended.[5] Such oversights may be particularly common because 'informativeness' is a word whose basic sense any competent English speaker can deduce; however, like many relatively familiar terms, it requires explication when used specifically as a technical term. Though the inconsistency of definitions can present problems, it also has the advantage of allowing for characterizations customized to particular work. One may, therefore, take advantage of the flexibility of the concept while avoiding confusion by presenting a clear characterization of informativeness along with discussion of how that characterization relates to previous accounts of informativeness; this is what I endeavor to do in this work.

In my work on adjectival past participles, I found that informativeness was an explanatory factor in the relative distribution of negated and non-negated participles in attributive position. I thus required a characterization of informativeness, and a way of assessing it; finding inspiration but no fully satisfactory accounts in the work of others, I developed my own characterization, which I present here. This is intended not to replace previous accounts of informativeness, but to elaborate on them in a way that has not previously been done in order to generate an account that is useful for purposes beyond those to which informativeness has thus far been applied. In particular, I apply informativeness to the purpose of explaining the relative frequency of adjectival past participles, though I maintain that my characterization of informativeness may be useful for other objectives as well. In brief, I propose that a phrase or utterance[6] is informative to the extent that it increases the cognitive availability of information not already fully available via category membership or context, and thereby reduces the number and/or likelihood of possibilities for the event, state, or entity to which it refers. I do not require an item to be of a propositional form in order to have the property of informativeness, though I do require that information be of a propositional form; and I treat cognitive availability as the ease with which something may be brought into conscious thought. I define context broadly to include one's cognitive environment, noting that aspects of

---

5 E.g., Wang & Inhoff (2010) in cognitive neuroscience and psychology, and van Leeuwen et al. (2009) in neuropsychology.
[6] 'Utterance' here and throughout this work is intended to any conversational turn or constituent thereof, regardless of modality; an utterance may have propositional or non-propositional form.

one's physical and cultural environment may be part of one's cognitive environment.

Although my characterization represents a development in the area of linguistic informativeness, the primary contributions of this work lie in its focus on information provided by category membership and its application of information and informativeness so conceived to the study of the distribution of adjectival past participles. In the simplest terms, information provided by category membership is information that one has about an item, based on prior knowledge of a category, when one knows the category to which an item belongs. The cognitive availability of such information is a function of how closely it is associated with the category in question, so that information about attributes necessary for category inclusion is fully available. In order to provide a structure for such knowledge and a way of assessing how it contributes to informativeness, I propose a model in which categories can be based on qualia or linguistic items and have gradient (or more- and less-central) members whose centrality represents cognitive availability. The portion of informativeness related to category membership, which I call CM-informativeness, can then be assessed via the centrality of a category member: when an utterance associates a linguistic item on which a category is based with some member of that category, the utterance is more informative if the member is relatively peripheral than if the member is relatively central. This relates to the fact that information that is less cognitively available prior to an utterance is that which, by being made fully available, has its cognitive availability most increased; one might think of this as the information that would be found the most "surprising."

The mechanisms described above are used to explain the distribution of negated and non-negated adjectival past participles in attributive position. An adjectival past participle occurs more often in non-negated form if its most common noun collocates are those with categories in which the participle is a peripheral member; this is because, in such a case, the negated version of the participle can be thought of as a central member. As stated above, an utterance associating a category base with a peripheral member is more informative than one associating a category base with a central member, so phrases with the non-negated form of the participle in question tend to be more informative than those with the negated form. Because, ceteris paribus, more informative utterances are expected

4

to occur more often than their less informative counterparts, in this case phrases with the non-negated form occur more often, leading to higher frequency of that form than of the negated form. In contrast, if an adjectival past participle's most common collocates are those with categories in which the participle is a peripheral member, the reverse occurs: phrases using the negated form of the participle tend to be more informative and occur more often than those with the non-negated form, leading to greater frequency of the negated form.

In chapter 2, I give an account of how the concepts of information and informativeness have been used in some influential theories. I begin in 2.1 with theories not associated with linguistics, recounting the general principles and early development of information theory, followed by the views of Dretske (1981) as a representative of the philosophical perspective on information and informativeness. Within linguistics, I discuss Grice's (1957-1991) work in some detail in 2.2.1, since it is the foundational work on implicature, the most prominent area in linguistics in which the concept of informativeness is used. Along with this I outline Horn's (1972) work on scalar implicature. I then briefly review the work of Horn (1984) in 2.2.2 and Levinson (1987; 2000) in 2.2.3, both of whom propose revisions to Grice's system. Finally, in 2.2.4 I provide a somewhat detailed account of relevance theory as developed by Sperber & Wilson (1995), as this theory shares a number of resemblances to my own proposals (not coincidentally – it was a major influence on my work).

Chapter 3 contains, in 3.1, a summary of the study on adjectival past participles that led to my formulation of an account of informativeness. As part of that account, I provide my initial characterization in 3.1.6. Then, in 3.2, I give a description of other factors that affect the relationship between informativeness and relative frequency of phrases of the type studied. These include the fact that more informative phrases are likely to be applicable less often, the influence of conventionality, and clashes in CM-informativeness. Following that, I provide a more current revision and expansion of my characterization of informativeness in 3.3, revising my stance on propositionality and, in particular, propositional form (it is no longer required of an informative item, though is required of information), clarifying what is meant by 'cognitive availability,' discussing

why category membership-related informativeness is considered to be possessed by the recipient prior to an utterance, and proposing my view of lexical items as having many qualia with a range of applicability.

Next, in chapter 4, I detail the category-based model described above. First, I discuss category attributes in 4.1, noting that categories have graded structure, are individualized, and can be common or ad hoc. Then, in 4.2, I show that a category can be based on a linguistic item and have qualia as members; for the sake of simplicity I examine only telic quale as members, and nominals as bases. A category can also be based on a quale, as I show in 4.3, and have linguistic items as members; again I confine my discussion to nominals and telic qualia, with the understanding that what is shown can be generalize to other types of linguistic items and qualia. For either category type, the CM-informativeness of a pairing between a category base and a member (or a linguistic item derived from base or member) is higher when the member is more peripheral, and lower when the member is more central. Next, in 4.4, I discuss interactions between categories of similar types. I then provide a generalized account of the types of assessments and predictions that the model allows for. Finally, in 4.5, I discuss differences between my model and other theories' treatments of informativeness; the most prominent of these is my focus on CM-informativeness, which is what the model assesses.

I conclude in chapter 5 by first reiterating the motivation for my work: to expand upon current characterizations of informativeness by providing an account that serves to explain the distribution of adjectival past participles and that has the potential to be useful for other purposes. Following that, in 5.2, is a review of what I have proposed and how this contributes to the field. In 5.3 I note some limitations of my work, and this leads to a discuss of future directions for research in 5.4. I provide concluding remarks in 5.5.

Chapter 2. Background: previous characterizations of informativeness

2.0 Introduction

In this chapter I discuss some background on the concepts of information and informativeness, particularly in the contexts of information theory and philosophy in 2.1 and several linguistic theories in 2.2. In the field of linguistics I focus first on Gricean implicature, which, as a theory proposed by a philosopher and used heavily in linguistics, can be thought of as bridging the gap between philosophy and linguistics. Following that, I discuss the principles of Horn (1984) and Levinson (1987; 2000), both of whom based their work on that of Grice (1975).[7] Finally, I examine relevance theory, which also grew partly out of work on implicature but makes use of some concepts from information theory as well.

2.1 Information and informativeness outside of linguistics

2.1.1 Information theory

Information theory (IT) is used primarily in mathematics, computer science, electrical engineering, and other fields. It operates under what is generally known as the code model, in which communication is viewed as a process of encoding followed by decoding. The model consists of the following components shown in (1) below (Ash 1965; Shannon 1948).

(1) a. Source: the origin of the message to be communicated.

     b. Transmitter or encoder: takes the message from its original form to that of the code.

     c. Channel: accepts input (coded message) in the form of a signal from the encoder; provides signal output to the decoder; is susceptible to noise, or interference that causes output to differ from input.

     d. Receiver or decoder: takes the coded output from the channel and renders it usable to the final recipient of the message.

     e. Destination: the final recipient of the message.

These components are illustrated below.

---

[7] 1975 is the year of publication in a book; the work is based on a speech given in 1957 and transcripts had been circulating since then, hence Horn (1972) being able to address topics treated in what is here referred to as Grice (1975).
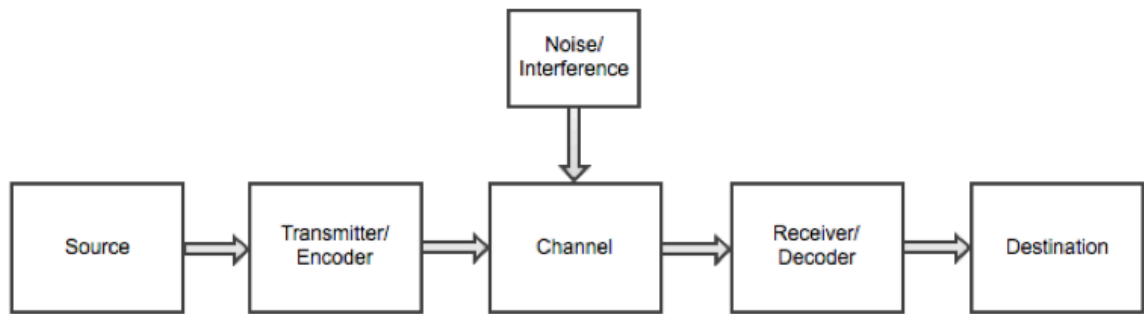
Figure 1: The code model of communication

One can see in figure (1) the path that a message takes from source to destination. IT was initially designed to deal only with discrete sources, or those that "generate the message, symbol by symbol" and "choose[s] successive symbols according to certain probabilities depending, in general, on preceding choices as well as the particular symbols in question" (Shannon 1948:5). However, it is no longer limited in this way. It is important to note that while the code model has been (and in some cases continues to be) used as a model for human linguistic communication, it applies to other types of communication as well. The source and encoder could be a speaker in a discourse, and the decoder and destination a hearer, but they need not be. This model has frequently been discussed (e.g., by Shannon 1948) with respect to radio, telephone, and television signal transmission, the types of transmissions for which it was originally developed, as well as transmission of other types of messages. While Shannon does address the information content of language specifically in his seminal 1948 work, it is with reference to written, rather than spoken, language.

The principal objectives of IT were originally to construct a mathematical version of the model described in (1) and shown in figure (1), and thereby to "provide methods for studying and discussing [signal distortion and interfering noise]" with the aim of maximally reducing these while minimally increasing the quantity of signal transmitted (Young 1971:7). Another way this goal is framed is, with a signal of set quantity, to maximally reduce distortion and noise while minimally reducing the amount of information actually transmitted via the signal; in other words, to send the greatest

amount of information possible, as clearly as possible, given the limitations of a signal and channel (Ash 1965). It is important to note that this goal is quite different from those of subsequent theories to be discussed. One of IT's important tenets is that there is a calculable value, the channel capacity, to which the amount of information transmitted per unit of time must be decreased in order to reduce noise and distortion to an arbitrarily low value and thus achieve "arbitrarily high reliability" of the transmission (Ash 1965:3). Channel capacity is measured in binary digits, or bits, per second. A bit is a unit representing the amount of information stored by a device in one of two possible positions, or the amount conveyed by eliminating one of two possibilities (Shannon 1948, Ash 1965). Here we see the first indication of informativeness having to do with the reduction of possibilities, as discussed in 2.0.

Maximizing the amount of information transmitted by a minimal signal necessarily involves defining and quantifying information, but despite its name, the measurement and especially the nature of information are less than central aspects of information theory. Sayre identifies a major "difference between [IT] and semantics" as "the difference between the study of the conditions for the communication of any message whatsoever and the study of the content of particular messages" (qtd in Dretske 1981:47). Dretske argues that IT may be "more properly viewed as a theory of *signal transmission*…. A genuine theory of information would be a theory about the *content* of our messages, not a theory about the form in which this content is embodied" (1981:40). On the topic of amounts of information, Ash notes that measuring the transmission rate of information is critical for the notion of channel capacity, and "...if we could pin down the notion of uncertainty, we would be able to measure precisely the transfer of information," which is defined in IT as that which removes uncertainty (1965:4). The amount of information associated with an event or state correlates with the reduction in uncertainty caused by that event or state, or how many possible outcomes were eliminated by it (Dretske 1981). Ash proposes a theorem for uncertainty based on the probabilities of all possible outcomes for a situation (Ash 1965: 8)[8].

---

8 Note that a very similar explanation and function appear in Shannon (1948:11).

Ash notes that uncertainty is independent of an outcome's possible values and "anything else except the probabilities associated with those values" (1965:11). He does not offer an assessment of whether this limitation is good or bad, but instead remarks, as an example, that the uncertainty of a unbiased coin toss "is not changed by adding the condition that the experimenter will be shot if the coin comes up tails" (1965:11). What Ash presumably means to convey is that, since in this case the probabilities referred to above are the probabilities of a heads or tails toss, something extraneous like the experimenter's well-being is completely irrelevant to the uncertainty of the situation and thus to the type or amount of information provided by a particular outcome. Though Ash does not, one could allow for some sort of modification to either the uncertainty equation or the calculation of probability values to permit the inclusion of experimenter-shooting probability, which would, in this scenario, be 1 for tails and 0 for heads, assuming that under this condition a coin flip of heads means the experimenter will *not* be shot. However, the same amount of uncertainty (regarding both the outcome of the coin flip itself as well as the fate of the experimenter) is removed by both heads and tails scenarios. Since information is viewed as removing uncertainty, this is equivalent to saying that both outcomes are equally informative.

Compare, however, a scenario in which an outcome of tails means that the experimenter will be shot, while an outcome of heads gives no clue about the experimenter's fate – he or she may or may not be shot. In this case, a crucial difference in assessment between IT and other theories discussed later in this chapter rests on the context in which uncertainty, information, and relevance are assessed. Following Ash's intimation that a set of probabilities must be for a *single* random variable, that is heads vs. tails, this situation is exactly the same as the one discussed above, in which heads meant safety for the experimenter. This is because, within the context of coin outcome *only*, to which probabilities are confined by Ash's calculations, the amount of uncertainty reduced by each outcome, and thus the amount of information provided, is identical to that of other outcomes in both scenarios. However, if one allowed for the uncertainty regarding the experimenter's fate to somehow be incorporated into the function as well, expanding the context in which outcomes are assessed, the same would not be true. Then, a tails

outcome completely removes all uncertainty about both the coin (tails) and the experimenter (shot), while a heads outcome removes all uncertainty about only the coin (heads) while all uncertainty remains about the experimenter. Within this larger context, the tails outcome clearly removes more uncertainty and is thus more informative, but this is not the type of context that IT's calculations are set up for; indeed IT basically ignores context, focusing instead on one discrete variable at a time.

In quantifying information, IT operates under the assumption that "any message which has a high probability of occurrence conveys little information," or in other words, removes little uncertainty (Young 1971:4). Thus the least likely messages convey the most information, and a message that is certain to occur conveys no information at all. The highest level of uncertainty, and thus the maximum information content, occurs when all possible symbols, messages, or outcomes have equal probabilities of occurring (Young 1971, Ash 1965).

IT does distinguish between the information content of a message and the amount of information actually received, by considering only what might be called 'new' information or content: "In order to obtain the actual information per symbol received, the total amount of the prior information per symbol possessed by the recipient must be subtracted from the information per symbol which the recipient would have received if he did not already have the prior information" (Young 1971:30). To apply this principle to human linguistic communication, one might imagine eliminating the phrase "per symbol" above. Thus the total amount of information in a message minus the amount that is prior knowledge of the recipient equals the total amount of information received from the message. If all the information in a message is also prior knowledge of the recipient, no information is received.

Although in assessing total information received it does take into account information that the recipient already has, IT does not place any importance on how the information in the message is conveyed. Thus, ceteris paribus, a message is considered equally informative whether it is delivered via gesture, Morse code, elaborate circumlocution, a simple statement in any language, or any other (ultimately successful) means (Young 1971). Intuitively, these are not all equally useful or appropriate ways to convey a

11

message under most circumstances, but they are considered informationally identical. On a similar point, Dretske argues that what the message recipient knows about the reliability of the signal or channel is irrelevant: "as long as the signal *is* reliable, whether or not it is *known* to be reliable… the amount of information reaching the receiver about the source… equals… the amount of information generated at the source" (1981:81).

Thus far in this section information has been discussed without reference to whether it is propositional or non-propositional, but it was noted earlier that information and informativeness are terms often used in discussions of propositions. Information theory is somewhat atypical in that the basics of the theory essentially conflate propositionality and non-propositionality, noting that information may be supplied by anything that "produces a message or sequence of messages to be communicated to the receiving terminal" (Shannon 1948: 2). A message was defined as a sequence of symbols (which may or may not amount to something propositional), any of various types of functions, or any other discrete source (as described with reference to figure 1). In addition to bypassing the issue of propositionality, or perhaps because of it, the relationship between information as defined by IT and truth remains unclear – must (propositional) information be true, or do propositions of unknown or false truth value qualify as information? Certainly a false message may be transmitted like any other, and since IT is concerned mainly with transmission rates and form over content, one might expect that the theory would allow for false information transmitted in a decodable manner.

Over time, information theory as a field has broadened significantly, so that what was initially quite theoretical and intended to apply to a specific type of communication signal is now used for a much greater variety of applications. The evolution of information theory was guided in particular by its use in radio, military communications, coding, the space program, microprocessing, modems, and digital wireless communication (Aftab et al 2001). Today, IT is used by researchers whose work deals with information in a number of ways, including processing, storage, transmission, and use, in a variety of fields, including electrical and computer engineering, computer science, mathematics, biology, theology, cognitive sciences, and many others.

2.1.2 Philosophy

In addition to outlining information theory, Dretske (1981) provides one of the most complete and recent accounts of information and its transmission from a philosophical viewpoint, so his views will be discussed here as representative of a philosophical perspective on information and informativeness. His stated goal is, as one may expect, very different from that of IT; it also, he notes, may be "too ambitious" (Dretske 1981:viii). That goal is to understand "how meaning can evolve, how genuine cognitive systems… can develop out of lower-order, purely physical, information-processing mechanisms" (Dretske 1981:vii). Dretske draws a distinction between meaning and information, proposing that one may consider the latter "as an objective commodity, something whose generation, transmission, and reception do not require or in any way presuppose interpretive processes" (1981:vii). This definition is quite similar to the information theoretical notion of the information content of a message, as distinct from the information actually received as the result of a message. Dretske notes that information is distinct from value, importance, truth, and knowledge; the fact that information is unaffected by the value a recipient may ascribe to it is one way in which information differs from meaning (1981:41,72). Despite the distinction between truth and information, Dretske writes that a message "may *have* a meaning but… *carries* information. What information a signal carries is what it is capable of "telling" us, telling us *truly*, about another state of affairs" (1981:44).

It is clear from the quote above that Dretske views truth as a qualifying characteristic of information: "*false* information and *mis*-information are not kinds of information" (1981:45). He notes that a lie transmits no information, or at least "no information of the kind I purport to be giving… you may nonetheless get information, information about *me* (I am lying), from what I say, but you will not get the information corresponding to the conventional meaning of what I say" (1981:44). It is interesting to note at this point Dretske's use of the word "conventional;" while he does not elaborate on this characterization, it seems to suggest an analysis that focuses away from inferential meaning, in sharp contrast with some of the theories to be examined later. Dretske defines a causal relationship between information and knowledge; knowledge "that s is F" is a "belief that s is F [which] is caused… by the information that s is F" (1981:86).

He emphasizes that, by this definition, "for a person's belief to qualify as knowledge, there must not only *be* evidence to support it, the belief must be *based on* that evidence" (1981:91). Since information is by definition true, and information must be the cause of the knowledge belief above, it follows that knowledge must also be true.

Dretske does not explicitly address whether information under his account is required to be propositional. However, his selection of the word 'telling' in the quote above on the difference between meaning and carried information, along with his repeated use of formulations such as "the information that *s* is *F*" (e.g., 1981:86), strongly suggests that propositionality is in fact a necessary or defining characteristic of information as he sees it. This appears to be a requirement that is shared with Dretske's notion of meaning: meaning is defined as being manufactured from information, and if the raw material is propositional, the product would most likely be as well (1981).

As with information theory, philosophical views of information transmission tend to take into account what information the recipient already has. However, there is not as clear a distinction between information carried by a message and information gained by a recipient of that message. Dennett (1969) notes that received information is difficult to quantify because "the information received by people when they are spoken to depends on what they already know" (qtd in Dretske 1981:79). Dretske, while agreeing with the quoted statement, appears to conflate the information content of a message and information received by the addressee, instead drawing a distinction between "the amount of information generated by the occurrence of an event and the amount of information a signal carries about that event," which he says can be compared using information theoretic formulae (1981:54). The information generated by an event or state of affairs is that which "a suitably placed observer" could gain from witnessing the event or state (Dretske 1981:45). He argues that the information contained in a signal is precisely quantifiable for a particular recipient by examining how much of the signal's information is not in the recipient's prior knowledge. Further evidence for the conflation of information carried and information received can be seen in Dretske's assertion that the relevance of prior knowledge to information transmission has only to do with any effect it has on "the amount of information generated at the source by the existence of a specific

14

state of affairs" (1981:81). He also notes, however, that this relativization of information is unimportant when all parties involved share the same relevant prior knowledge, and argues that this is a common circumstance (1981:80). This aligns with Wilson and Sperber's (2012) suggestion, noted later in this chapter, that a speaker generally knows the relevant prior information held by their addressee.

Concerning the transmission of information, Dretske recognizes a principle of transitivity and a "communication condition" that states that "if a signal carries the information that *s* is *F*, then it must be the case that… the signal carries as much info about *s* as would be generated by *s*'s being *F*," that it is in fact true that "*s* is *F*," and that the signal actually does carry the information "generated by *s*'s being *F*" (1981:57, 63-64). The first part of this condition is something like a principle of conservation of mass, but for information quantities instead, while the second reiterates the postulate that information must be true. The last part, though seemingly vacuous, prevents even perfect correlations from being identified as informational relations: "even if… whatever is *F* is *G* and vice versa… this does not mean that there is information in *s*'s being *F* about *s*'s being *G* (or vice versa)" (Dretske 1981:73).

Though this section on informativeness outside of linguistics is by no means exhaustive, information theory and the views of Dretske represent two of the most important areas firmly outside of linguistics in which informativeness is used. I now move on to examining the concept of informativeness in areas more clearly related to linguistics. The next section is arranged chronologically, but this has the welcome effect of making Grice the next scholar whose work I examine. As a philosopher of language, he is an excellent person with whom to make the transition from discussing informativeness outside of linguistics to discussing it within linguistics.

2.2 Information and informativeness within linguistics

2.2.1 Gricean implicature

Grice's theory of conversational implicature is one of the best-known theories within linguistics that uses the concept of informativeness. Due to later development by Horn (1972), conversational implicature is also an area of linguistics in which informativeness has a relatively accepted, precise definition – the entailment-based definition mentioned

15

in 2.0, which is discussed more later. Grice's work in this area stems from the general acceptance among logicians that there is a "divergence in meaning between… some… formal devices" of logic, such as ¬, ∧, ∨, and ∃x, and corresponding linguistic expressions like *not*, *and*, *or*, and "*some* (or *at least one*)" (Grice 1975:41). Grice intended to show with his analysis that such a divergence does not actually exist, and that "conditions governing conversation" account for the apparent differences in meaning (1975:43). In other words, the differences between formal logical devices and their corresponding linguistic expressions lie not in meaning, but in possibilities of interpretation. In addition, his theories account for the fact that, even outside the use of logical operators, "what a speaker intends to communicate is characteristically far richer than what she directly expresses; linguistic meaning radically underdetermines the message conveyed and understood" (Horn 2004:3). Accounting for this fact is the basic motivation behind, and goal of, all the theories discussed in the remainder of this chapter. It can thus be seen that, based on their goals, the theories discussed throughout this chapter form three groups: IT, philosophy, and the remaining more linguistic theories.

2.2.1.1 The cooperative principle and conversational implicature

Grice first introduced what he called "the Cooperative Principle," "a rough general principle" that, under usual circumstances, interlocutors are expected to adhere to.

(2) The Cooperative Principle

> Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

(Grice 1975:45)

The purpose of this principle was to capture the fact that a conversation generally does not comprise "a succession of disconnected remarks," and usually involves at least some amount of "cooperative effort" with interlocutors agreed, at a minimum, on "a mutually accepted direction" (Grice 1975:45). To go along with the cooperative principle, Grice proposed four conversational maxims, shown in (3).[9] It is important to note that the

---

[9] Green (1996) notes that the maxims are explicitly *not* intended to constitute the cooperative principle, as some have mistakenly thought, but to accompany it.

maxims are intended not as rules per se, but as "unstated assumptions we have in conversations" or "default settings" of which interlocutors are mutually aware (Yule 1996:37; Horn 2004:8).

(3) Conversational maxims

    a. Quantity: Make your contribution as informative as is required (for the current purposes of the exchange). Do not make your contribution more informative than is required.

    b. Quality: Try to make your contribution one that is true:

        Do not say what you believe to be false.

        Do not say that for which you lack adequate evidence.

  c. Relation[10]: Be relevant.

    d. Manner: Be perspicuous:

        Avoid obscurity of expression.

        Avoid ambiguity.

        Be brief (avoid unnecessary prolixity).

        Be orderly.

(Grice 1975:45-6)

The cooperative principle and the maxims capture the fact that interlocutors generally assume that those with whom they are interacting are speaking appropriately[11] – namely, in a truthful and relevant manner, with the right amount of information, and without deliberate obfuscation. When necessary, a speaker may choose to partially or completely opt out of a maxim through the use of hedges such as *I'm not sure, but maybe…* (hedging quality) or *Oh, by the way,…* (hedging relation) (Green 1996; Yule 1996). However, in the absence of evidence to the contrary, a presumption is made that the speaker is obeying all maxims,[12] and a great deal of information is obtained as a result of this presumption, via what Grice calls conversational implicatures.

---

[10] Also known as "relevance."

[11] What is considered appropriate varies cross-culturally; the maxims themselves or how they are obeyed may need to be adjusted to fit other communicative cultures (Mey 2001).

[12] Grice (1978) distinguishes this from separate presumptions that the speaker is obeying each maxim individually.

An implicature, according to Grice, is an inference made by a hearer about what a speaker has "implied, suggested, meant, etc." (Grice 1975:43). Though Grice is less than clear on the point, Levinson asserts that Grice means to classify as implicatures only those inferences which are "intended to be recognized as having been intended," thus excluding from classification as implicatures those inferences that a hearer may make on the basis of errors or unintentional cues by the speaker (1983:101). Information that is implicated is distinct from what is actually said. What is said comes from the conventional meaning of the words in an utterance, along with the grammatical structure of the utterance and additional information like referents, disambiguations of words and phrases, speaker, and time and place of utterance (Grice 1975; Horn 2004).[13] A conventional implicature is what is indicated by the conventional meanings of the words (but, again, is not actually said). For example, Grice provides the utterance "*He is an Englishman; he is, therefore, brave*" and notes that the speaker of such an utterance will not have "said (in the favored sense) that it follows from his being an Englishman that he is brave," but rather will have "certainly indicated, and so implicated, that this is so" (1975:44-45). The distinction here between what is said and what is implicated rests on Grice's (1975) claim that, should it turn out that "he" is in fact an Englishman and brave, but that his bravery is not a consequence of his Englishness, the utterance would not be false. In later works, Grice makes similar points about the use of *but*, *on the other hand*, and *so* (Carston 2002).

In contrast, a conversational implicature (hereafter referred to as just an "implicature," as conventional implicatures will not be further discussed) is what is conveyed not by what is said or what conventional meanings indicate, but by "certain general features of discourse," namely, those in (2) and (3) (Grice 1975:45). Unlike conventional implicatures, conversational implicatures can be eliminated or strengthened (Yule 1996; Horn 2004); this is discussed in section 2.2.1.4.3. Grice further distinguishes between generalized and particularized conversational implicatures, the former "aris[ing] in such a broad range of contexts that it will not seem to depend on the details of the context too much" while an instance of the latter relies on "special features" of a

---

[13] For a more formal treatment of "saying," see Récanati (1991).

particular context and does not normally arise with every use of the linguistic forms that accompany it (Geurts 2010:17; Grice 1975:56; Horn 2004). Many scholars (such as Neale 1992 and Geurts 2010) treat this distinction as unimportant, and it will not be further discussed here. However, many examples given will be what Grice would consider particularized.

An implicature can arise either from the presumption that a maxim is being obeyed, or when a maxim seems to have been violated. Violation may result from a "clash" between maxims – a situation in which two maxims are in conflict, and in order to obey one maxim, another must be violated, and vice versa (Grice 1975:49; Green 1996). Violation may also occur either in order to deceive or in order to convey a message other than what is strictly said (in the sense of "said" discussed above). Grice refers to the latter type of violation, or "blatantly fail[ing] to fulfill" a maxim, as flouting a maxim (1975:49); this is discussed in 2.2.1.3 and in parts of 2.2.1.4.

2.2.1.2 Implicatures that arise when maxims are obeyed

Implicatures caused by the presumption that the speaker is obeying a maxim will hereafter be referred to as "standard implicatures," following the terminology of Levinson (1983). They routinely arise in a talk exchange in the absence of mitigating factors, or "wherever features of the context do not actually block them" (Levinson 1983:108). A hearer's assumption of cooperation on the part of a speaker, together with an interaction in which the speaker appears to obey the conversational maxims, enables the hearer to easily make a number of inferences regarding the truth, relevance, informational appropriateness, and clarity of the speaker's utterance. Due to the maxim of quality, the hearer infers that what the speaker says is true, or at least the speaker thinks this is the case and has adequate evidence to think so. The maxim of relation allows the hearer to assume that what the speaker says is relevant to the point of the discourse. Because of the maxim of manner, the hearer accepts that the speaker's way of speaking is as clear, cogent, and appropriate to the hearer's characteristics as possible. And the maxim of quantity leads the hearer to understand that the information the speaker provides is enough, but not too much, for the point of the discourse. Such inferences are frequently so unremarkable and automatic that they may seem barely worthy of

discussion, but it is precisely because they are usually given so little thought that it is appropriate to explicitly outline them.

Some example standard implicatures can be seen to arise from the exchange shown in (4).

(4)   A: What time is it?

        B: It's ten fifty-one.

Here B intends to convey, and A is able to assume, based on the maxim of quality, that it is actually 10:51, or at least that B thinks it is and has good reason to do so; for example, B may have just looked at a clock, not known to be inaccurate, that said 10:51. Unless there is any context to suggest otherwise, A can further conclude, by virtue of the maxim of relevance, that 10:51 is the time in A's current location, since this is the most relevant time for A to know.[14] Standard implicatures resulting from the presumed obeying of the quantity and manner maxims are easier to see with the exchange in (5):

(5)   A1: I was really busy this morning.

        B: What did you do?

        A2: I walked the dog, took a shower, ate breakfast, finished my grant proposal, and bought and mailed a birthday present for my sister.

B is justified in concluding that A meant to convey that he did the things in A2 in the order stated, and that, absent any context indicating otherwise, none of the descriptions given are euphemisms or obscure phrases used to indicate activities other than those that are immediately called to mind (e.g., A does not mean that before his shower he engaged in the yo-yo trick known as "walk the dog"). Both of these conclusions are motivated by B's assumption that A is obeying the maxim of manner in A2. Further, on the basis of the presumption that A is obeying the maxim of quantity, B may infer that the activities A mentioned were the main or most important things that he did that morning. However, because the quantity maxim forbids both overinformativeness and underinformativeness,

---

[14] Examples of context that might suggest otherwise: B is in another time zone and has mentioned needing to get to bed (so B's local time would be relevant); or A needs to call someone in another location and B has warned A not to do it right now because it may awaken them (so the time in the location of the person A needs to call would be relevant).

B may also surmise that the list of activities is not exhaustive, and that A did other things not worth mentioning, such as wake up, breathe, and get dressed.

As shown in (4) and (5), implicatures can arise without the violation of a maxim, simply as a result of a hearer's supposition that a speaker is being cooperative, and this supposition causes standard implicatures of several types. Such implicatures arise as a result of the cooperative principle together with the conversational maxims, and are foundational for the creation of any further (non-standard) implicatures.

It is worth noting that standard implicatures rely a great deal on calculations or decisions made by the speaker: what constitutes adequate evidence of truth; what the point of the discourse is; what the hearer's characteristics are and what is appropriate to someone with these characteristics; what information the speaker already has (and therefore does not need, even if it is necessary for the point of the discourse). Any difference of opinion between speaker and hearer on any of these points can result in the speaker unwittingly violating a maxim and motivating the hearer to make other inferences which, being unintentionally triggered by the speaker, would not be classified as implicatures.

2.2.1.3 Implicatures that arise when maxims are violated

A speaker who has adequate and accurate knowledge of what is appropriate/expected can choose to say something conspicuously inappropriate or unexpected in order to create an implicature by using the cooperative principle and one or more maxims to send a message distinct from the conventional meaning of his/her utterance: "we can make a blatant show of breaking one of the maxims… in order to lead the addressee to look for a covert, implied meaning" (Leech & Thomas 1988:15-16). A hearer who perceives an apparent violation of one or more maxims is presented with a dilemma: a maxim seems to have been violated, but the hearer generally maintains a supposition that the speaker is following the maxims. Therefore the hearer's reaction to this situation is to search for a way in which the speaker may not actually be violating a maxim. In Grice's (1975) explanation, the hearer may have something like the thought process shown in (6):

(6)   The speaker's saying of x seems to violate one or more maxims.

        If the speaker, by saying x, intends me to understand y, then the speaker is not

violating a maxim, and my supposition about his cooperativeness may remain. [15]

The speaker may reasonably think y.

The speaker may also reasonably think that I have the ability to discern that y is what he intends me to understand, though he has said x. [16]

As a result of this thought process, the hearer will suppose that the speaker did, in fact, intend y to be understood, and will thus understand y. Then y is an implicature caused by the apparent violation of a maxim, and the speaker has conversationally implicated y. This is the basic process by which intentional apparent violation, or flouting, of a maxim can cause implicature. As described by Levinson, this process is "based on the remarkable robustness of the assumption of co-operation: if someone drastically and drastically deviates from maxim-type behavior, then his utterances are still read as underlyingly co-operative if this is at all possible;" the speaker's apparent violation of a maxim thus "force[s] the hearer to do extensive inferencing to some set of propositions, such that if the speaker can be assumed to be conveying these then at least the over-arching co-operative principle would be sustained" (1983:109).

Examples of implicatures caused by flouting can be seen with reference to the exchanges shown in (7), (8), and (9). In (7), the maxim of quality is flouted by B.

(7)   A and B live 1000 miles apart. B has many obligations at home, but is coming to visit A in June, and they are discussing the duration of the trip.

A: I wish you could stay longer! You know the weather here will be much better than the weather there in late summer, so you could just not even go home until October.

B: Oh okay, sure, I'll just tell everyone here that they'll have to live without me for five months!

In this exchange, A has expressed a suggestion that is plainly infeasible and unrealistic: that B visit for five months. B's response is sarcastic and untrue – B does not intend to

---

[15] It is arguable that a speaker who says x in order to convey y is necessarily violating the maxim of manner; one may consider this a clash between the manner maxim and the reason that the speaker has for choosing flouting, where the speaker has opted to prioritize the reason for flouting over obeying the manner maxim.

[16] In later work, Grice uses the term "M-intend" to capture intent by a speaker to produce an effect in a hearer "by means of [the hearer's] recognition of that intention" (1991:68).

tell those at home to do without her for five months – and this is intentionally made clear to A. In this case, B's intended implicature, and the message that A receives, is that B thinks A was silly to even propose a five-month visit.

In (8), A flouts the maxim of manner.

(8)   B has lots of experience with young children, and A knows this. B has come to A's house to babysit A's child, who is currently very cranky.

        A: Somebody definitely needs to take an N-A-P [pronouncing the names of the letters].

By her utterance, A is engaging in ambiguity ("somebody") and obscurity of expression (by spelling rather than saying the word "nap"). Because A has blatantly violated the maxim of manner in these ways, B is able to infer that the child should not be made aware of A's assessment. A relies on B's experience with children to enable her to figure out that this is because, as is common with young, tired children, the child is irritable and is denying being tired, so hearing someone say that he should take a nap would only make him crankier.

In (9), the maxim of relation is flouted in order to more efficiently send a message that could have been conventionally said.

(9)   A and B are running errands, with B driving. They had planned to stop at the hardware store to replace a faucet fitting. Approaching the road that the hardware store is on, B has not activated a turn signal.

        A: Aren't you going to turn?

        B: I forgot the fitting.

Here B has flouted the maxim of relation by saying something that, by its conventional meaning alone, does not constitute an answer to A's question. However, A is able to infer that B is not going to turn. This is based on the fact that A presumes B intends to answer the question, and knows the reason for turning is to go to the hardware store, the reason for going to the hardware store is to replace the fitting, and in order to be sure that the correct fitting is purchased, they must have the old one with them. Thus flouting the maxim of relation allows B to provide a very concise but adequate answer to A's question.

2.2.1.4 Quantity implicature

While both types of implicature – standard and flouting – can be discussed with reference to each of the maxims and submaxims in (3), the focus here is on the maxim of quantity, which, in enjoining speakers to provide neither too little nor too much information, necessarily engages the issue of the measurement of informativeness. Interestingly, Grice neglected to define informativeness or a way of measuring it when he proposed his maxims of conversation; while informativeness is often measured in terms of entailment, this was not explicitly discussed with reference to implicature until Horn's (1972) dissertation, which is discussed later in this chapter. Regarding the amount of information an interlocutor provides, Grice wrote simply that a cooperative speaker should "make [his or her] contribution as informative as is required (for the current purposes of the exchange);" this was intended to classify as uncooperative both over- and under-informative contributions (1975:45). It is unclear whether, in Grice's view, something must be propositional in order to be informative; however, some of his discussion with respect to the quantity maxim suggests that propositionality is not necessary for informativeness. For example, Grice (1989) notes that using an expression of the form *an X* can be considered less informative than using a similar expression in which the owner of *X* is identified, e.g. *my X*. This would seem to indicate that non-propositional items can be informative, but it could also mean simply that a proposition using *an X* is less informative than one using a form such as *my X*. Further, Grice later states that "false information is not an inferior kind of information; it just is not information" (1989:371); this seems to indicate not only that information must be true, but that it must be propositional – since truth is a property of propositions. Then, assuming that to be informative something must contain information, informativeness would in fact be a property limited to propositions.

Grice is even less explicit on the subject of whether an addressee's previous knowledge should be considered when determining the informativeness of an utterance. However, he says nothing to suggest that previous knowledge should *not* be considered, and it seems reasonable to interpret the phrase "for the current purposes of the exchange" from the maxim of quantity as including consideration of one's partner in the exchange

and information he or she might already have (Grice 1975:45). While informativeness itself is not noted to be affected by other attributes of a message, such as how it is conveyed or how it relates to other information, some of these are clearly addressed by other maxims, particularly those of manner and relation.

Whether or not an utterance has the appropriate or expected level of informativeness for the discourse, implicatures may arise as a result of the level of informativeness. These can be classified into two categories: those relating to overinformativeness, and those relating to underinformativeness. Implicatures related to overinformativeness are the standard implicatures that arise from the presumption that a speaker is not being overinformative, as well as the implicatures that arise from a speaker flouting the quantity maxim by being intentionally apparently overinformative. Correspondingly, the implicatures related to underinformativeness are the standard ones that result from the hearer presuming that the speaker is not being underinformative, and the implicatures created when a speaker flouts the quantity maxim by being intentionally apparently underinformative.

2.2.1.4.1 Overinformativeness

As noted above, standard implicatures arise when a hearer presumes that a speaker is *not* being overinformative, or is not giving more information than is needed or appropriate. In particular, the hearer infers that all of the information being given is necessary and/or appropriate for the discourse. For example, if A asks B how to prepare a particular food, and B gives directions that consist of five steps, A will infer that B meant it is necessary to complete all five steps. Even when a speaker appears to be overinformative, the hearer will try to maintain his/her conception of the speaker as cooperative; this fact enables the speaker to be intentionally overinformative in order to trigger one or more non-standard implicatures.

Intentional apparent overinformativeness, or flouting of the quantity maxim, can be used to trigger implicatures with several different message types, shown below:

(10) a. The speaker wants the hearer to know the full extent of his/her knowledge.

b. The speaker thinks that the hearer should have more information than the hearer wants or expects.

c. The speaker wants to make the hearer think that the topic is more difficult or complicated than it actually is, by portraying that all the information given is necessary.

Most of these will be seen to directly contrast with implicature types caused by intentional apparent underinformativeness, given later in (23).

(10a) is a type of implicature that may arise in a situation in which the speaker is trying to impress the hearer, or perhaps dispel any notions the hearer may have about the speaker's ignorance. (11) shows an example of what would most likely be considered the former motivation:

(11)   A is a teacher speaking to his class; B is a student in the class.

A: What do we call the structure that bees live in?

B: Bees live in a hive, which is made of beeswax and has lots of cells, each with six sides. Each hive has a queen, who grows up in a special cell of the hive.

Here the student (B) has given the information A asked for, plus quite a bit more related information. Assuming that B is aware that this is more information than was necessary or expected, A and B's classmates are likely to infer that B wants them to be impressed with his knowledge of beehives and for this reason has chosen to share more of it than was necessary.

In the type of implicature in (10b), B again gives A more information than B knows is desired or expected, but this time it is because B thinks A should have this information – perhaps A needs the information but does not realize it, or A should (in B's opinion) want the information. An example in which this type of implicature might arise is below:

(12)   A is going to housesit for B, and B hands A a set of keys.

A: Which key is for the front door?

B: The red one, which you also have to use in order to lock the door. When you're locking or unlocking it, you need to pull the door towards you in order for the key to turn easily.

In this situation, B has given more information than A explicitly asked for or was likely expecting, but it is clear that B thinks this information is necessary for A, and wants A to realize that; knowing which key to use to unlock the door is not particularly useful if one

26

does not know the trick to get the key to actually work, or how to lock the door afterwards. A thus gets the implicature that this other information is also important, realizing that he needed it even though he did not know enough about the situation beforehand to explicitly expect it.

The type of implicature in (10c) contrasts with that in (10b) in that, in (10c), B gives extra information not because she thinks it is necessary, but because she wants A to think that it is necessary. A's intent here is thus quite similar to B's intent in (12), but with different motivation: in (12) B wanted A to think the information was necessary because it was, in fact, necessary, while in (13) A wants B to think the information is necessary despite the fact that it is not. Causing A to think that more information is necessary is a way to give him the impression that the topic under discussion is more difficult or complicated than he thinks it is (and than it actually is). B may desire to cause this impression in order to intimidate A or to dissuade him/her from some course of action, as in (13).

(13)   A is an amateur violinist who asks a luthier friend, B, for advice on repairing a bow.

B thinks that A is overestimating his precision woodworking abilities, and that the repair should be done by someone more qualified and with better equipment, but does not want to offend A by saying so directly.

A: So, my bow has a crack in the tip. It's a pretty small crack, and it was a cheap bow to begin with, so I thought I'd try to repair it myself and see how that goes. I've read a bit, but can you tell me the basic process you'd use?

B replies to A's request for advice with an extremely long and detailed set of instructions on how to repair the bow, with lots of description of warnings to be extremely careful.

B has almost certainly given A more information than he was expecting; B's hope in doing so is that A will get the implicature that B thinks the repair is beyond A's capabilities, and that A will allow B or someone else more trained in luthiery to do it.

In this section I have shown that implicatures arise from the assumption that speakers are not overinformative in usual circumstances. Implicatures also arise when a speaker

appears to be intentionally overinformative, as a result of the hearer inferring the reason for the speaker's apparent overinformativeness. In the next section, I discuss the same processes with respect to underinformativeness.

2.2.1.4.2 Underinformativeness

As with overinformativeness, standard implicatures result from a hearer's presumption that a speaker is not being underinformative, or is not providing less information than is needed or appropriate. One such type of standard implicature is that any information that was not provided was either not appropriate or was unnecessary. The second type of standard implicature caused by assuming a non-underinformative speaker is known as a scalar implicature, and involves the hearer's presumption that the information provided is as informative as is possible given the other maxims.

For example, if A tells B *I have two dogs*, B is able to infer that A means to say that she has exactly two dogs, and in fact does have exactly two – no more, and no fewer. This is based on two maxims. The maxim of quality enables B to assume that A is not lying,[17] and as such, she must have *at least* two dogs, since her statement would be false if she had only one or zero dogs. The maxim of quantity provides the motivation for B's assumption that A has *at most* two dogs, since, if she did have more than two dogs, A would be withholding information that it is appropriate to supply in this conversation (the information that she has one or more additional dogs).[18] Taken together, then, the quality-based assumption of *at least two* and the quantity-based assumption of *at most two* yield a combined assumption of *exactly two*.

The implicature that B makes in assuming that A has at most two dogs is an example of a scalar implicature; this type of implicature is one of the primary and most well-known areas of linguistics in which informativeness and its relative quantification are used. Scalar implicatures were first written about extensively by Horn (1972), who introduced the notion of scales on which linguistic items are ordered such that an item

---

[17] Note that all other implicatures are basically dependent on the assumption that the quality maxim is being obeyed; many, including Grice himself, have taken this position and thus given the quality maxim status as the most important of the maxims (Grice 1975; Levinson 1983; Horn 1984, 2004).

[18] If A happened to have five dogs, she has not actually *lied*: having five dogs, she also necessarily has four dogs, three dogs, two dogs, and one dog. The strength of the quantity implicature, however, is such that this type of information withholding is frequently perceived by the hearer as a lie, when in fact one may differentiate between lying and deception, and classify this withholding as the latter.

entails all items to its left but does not entail any of the items to its right.[19] The ordered linguistic items are, of course, not randomly selected: "scalemates" are taken to be relevant in some shared context (Geurts 2010). It is then said that the rightmost item that accurately describes a situation is the most informative in that situation, with informativeness effectively defined as the number of entailments regarding the relevant scale. According to Horn (1972), the use of each item on the scale implicates that none of the items to its right could accurately be used. Therefore asserting any item on the scale implicates the negation of all items to its right (Yule 1996). This effect arises because of Grice's injunction to "be as informative as is required" (Grice 1975:45).

A very simple scale of this type can be seen with the cardinal numbers; as noted in footnote 17, that there is (or that one has, etc.) a number n of items entails that there is also every number from one to n of that item, but it entails nothing regarding the possibility of there being more than n items. As Horn (1972:41) says, sentences containing numbers "assert lower-boundedness – at least n – and… may, depending on the context, implicate upper-boundedness – at most n – so that the number may be interpreted as denoting an exact quantity." Thus a "Horn scale" for the cardinal numbers is as follows, where $<<$ indicates one-way entailment and can be read 'is entailed by.'

(14) $1 << 2 << 3 << 4 << 5 << 6\dots$

Returning to A and her dogs, one can see again how B reached the conclusion that A has exactly two dogs: it is implicated by the quantity maxim that A is using the most informative quantity possible to describe her dog-owning situation, meaning that none of the numbers to the right of two on the scale must be available for use (i.e., to use them would not be truthful, so A must have no more than two dogs). Further, since no number below two entails two, A has asserted that she has at least two dogs, so in B's view she must have no fewer than two, since B assumes she is obeying the maxim of quality. Once again, the conclusion is reached that A has exactly two dogs, using the "general fact" that "scalar predicates [and quantifiers] are lower-bounded by assertion and upper-bounded by implicature" (Horn 1972:42). However, as shown above, this conclusion is relatively

---

[19] Some have proposed scales based on relationships other than entailment (e.g., Hirschberg 1985).

easy to reach without directly referencing a scale such as that in (14). It is with less obvious entailment situations that Horn scales become more useful.

One of the first scales proposed by Horn was for quantifiers, and is shown below:

(15) one << some / a few << several << many << half << most / a majority << all / every

<div align="right">Horn 1972:61</div>

This indicates that *several*, for instance, entails both *some* and *one*, and implicates some amount less than *many*, *half*, *most*, or *all* (Horn 1972). Thus, according to Horn's (1972) analysis, in a situation where (16a) is not true but (16b) is, (16c) is also true but less informative than (16b), and (16d) is even less informative than (16c), though still true.

(16)   a. There are many trees in the yard.

   b. There are several trees in the yard.

   c. There are some trees in the yard.

   d. There is one tree in the yard.

If one construes *some* as less than *several*, as Horn does, then certainly (16c) is true if (16b) is, though (16c) fails to fully describe the extent to which there are trees in the yard. Likewise, certainly there is one tree in the yard if there are several, but it would be pragmatically uncooperative to use (16d) when either (16b) or (16c) was true. Thus a hearer of (16d) will, under most circumstances, conclude that the yard in question probably does not contain enough trees for (16a), (16b), or (16c) to be accurate, i.e., there is exactly one tree in the yard.

Horn (1972) provides a number of other examples of items that operate on a scale with respect to implicature; in most of these cases, the scale is like those in (14) and (15), with one-way leftward entailment. Among predicates, scales can be seen for certain adjectives. One such example is the coolness scale below:

(17)   lukewarm << cool << cold   (Horn 1972:38,48)

Several things must be noted about this scale. First, it may seem counterintuitive to think of each item on the scale as entailing the items to its left; this difficulty, however, can be surmounted by thinking of these entailments in terms of the scale itself, rather than in absolute terms. That is, while one may find (18a) objectionable, (18b) seems to be acceptable because it references the scale in question.

<div align="center">30</div>

(18)   a. *Cold* entails *lukewarm*; if the water is cold, it must also be lukewarm.

        b. *Cold* entails *at least as cool as lukewarm*; if the water is cold, it must also be at

        least as cool as the temperature designated by *lukewarm*.

Second, one may wonder why there is a coolness scale (and, though not yet mentioned, a

corresponding warmness scale) instead of a single temperature scale as in (19a) or (19b).

(19)   a. * cold << cool << lukewarm << warm << hot

        b. * hot << warm << lukewarm << cool << cold

Such a scale cannot exist because of problems with the directionality of entailment.

Essentially, the right half of each scale above is correct: *cold* entails at least (as cool as)

cool, and *hot* entails at least (as warm as) *warm*, but none of the statements in (20), all of

which result from interpretation of the scales in (19), are acceptable.

(20)   a. *Cool* entails at least as warm as *cold*.

        b. *Cool* entails at least as cool as *cold*.

        c. *Warm* entails at least as warm as *hot*.

        d. *Warm* entails at least as cool as *hot*.

(20a) and (20b) are derived from the scale in (19a), and vary depending on whether one

construes said scale as one of warmness or coolness. (20c) and (20d) are similarly

derived from the scale in (19b). (20a), while not obviously false, is anomalous or

underinformative at best. If one accepts the somewhat problematic notion of talking

about the relative warmth of *cool* and *cold*, *cool* could in fact be said to entail "warmer

than *cold*." (20d) is comparable; while it is questionable to talk about the relative

coolness of *warm* and *hot* in the first place, if this is accepted one can say that *warm*

actually entails "cooler than *hot*." (20b) and (20c) are simply false: *cool* is in fact less

cool than *cold*, while *warm* is less warm than *hot*. From this it can be seen that neither of

the scales in (19) are possible. Further proof of this can be found by examining

implicatures, which can be felicitously contradicted, and assertions, which cannot. If

(19a) were correct, *cool* should implicate *not warm*, just as *some* implicates *not all*.

However, *cool* in fact asserts *not warm*, as can be seen by the fact that *not warm* cannot

be contradicted:

(21)   *It's cool; in fact it's warm.   (Horn 1972:39)

31

Similar results are found with implicatures predicted by the scale in (19b); *hot* does not implicate *not cool*, but rather asserts it. Thus there is ample evidence for separate scales of coolness and warmness.[20]

Scales based on other adjectives also have corresponding opposite scales. This existence of corresponding "positive" and "negative" scales, as Horn calls them, is extremely common among scalar predicates, and interacts interestingly with the markedness of questions and answers: a neutral question (i.e., one that gives no information about the expected answer) that requests information on a scale "will employ a relatively weak element on the positive scale" (Horn 1972:72). In contrast, if the answer is "expected to fall on the negative scale," the question is most likely to use a somewhat weak item from the negative scale. A question using a strong item from the negative scale likewise indicates expectation of a negative scale answer, but here the answer is expected to be a strong negative item, rather than simply any item from the negative scale. Similarly, a question that uses a strong positive item suggests that the asker expects an answer on the positive scale, and perhaps a strong positive. Consider again the scales of warmness and coolness; here the "positive" scale is that of warmness and the "negative" scale that of coolness. The following questions illustrate the effects Horn noted:

(22)   a. How warm is it?   It's a bit cold, pretty cool, very warm, extremely hot

         b. How cool is it?   It's a bit cold, pretty cool, # very warm, # extremely hot

         c. How cold is it?   It's a bit cold, (#) pretty cool, # very warm, # extremely hot

         d. How hot is it?   It's extremely hot, (#) very warm, # pretty cool, # a bit cold

Any item on the warmness or coolness scale is an acceptable answer to (22a); a sample of these is shown. With (22b), however, only answers on the coolness scale seem highly appropriate, while those on the warmness scale are somewhat anomalous. For (22c), the given answers go in order from best to worst, or least to most anomalous, indicating that a strong negative answer is the most expected. The same effect is shown in (22d); note that the answers, again given from best to worst, are in exactly the opposite of the order in (22c).

---

[20] A final note is in order to remark that *lukewarm* can appear on both scales (leftmost in both cases), depending upon speaker and context.

Of course, standard implicatures, such as scalar implicatures, are not the only ones related to underinformativeness. A speaker who flouts the maxim of quantity by intentionally giving what seems to be too little information is likely to trigger one (or more) of a few basic types of implicatures, shown in (23).

(23)   a. The speaker has no further (truthful, well-evidenced) information to give.

b. The speaker thinks that the hearer should not have any further information.

c. The speaker wants to make the hearer think that the topic is easier or simpler than it actually is, by portraying that the information given is all that is necessary.

d. The speaker wants the hearer to feel that the hearer's prior contribution(s) to the discourse was inappropriate.

(23a), (23b), and (23c) correspond (and contrast) roughly to (10a), (10b), and (10c), respectively: both (a) items concern the amount of the speaker's knowledge, (b)s involve the speaker's opinion about what information the hearer should have, and in both (c) items the speaker is attempting to deceive the hearer with regard to the simplicity or complexity of the topic of discussion. While one could imagine a (10d) that contrasts with (23d), such an implicature type is not used commonly enough to be discussed here.

(23a) represents a type of implicature likely to arise when a vague answer is given to a question that requires a more specific answer, as in (24):

(24)   A is supposed to drive B to a doctor's appointment, and on the morning of the appointment:

A: What time is your appointment?

B: Some time between ten and noon.

Clearly A needs more information in order to get B to the appointment on time, and there is no reason (given the relatively limited context described here) to suppose that B is purposely being uncooperative, so A is likely to get the implicature that B intends to convey that he does not currently recall the exact time of the appointment. Further, A may expect that B will, at some point relatively soon, find out the exact time and provide that information.

The type of implicature in (23b) contrasts with that in (23a) in that it involves the speaker having further information, but simply not wanting to share it. This may occur if

the speaker views the information in question as private (at least with respect to the hearer), or if the information is necessary for the hearer to accomplish something that the speaker wants to thwart for some reason. The former is probably the more common situation, and is illustrated below:

(25)   Strangers are making small talk while waiting for a bus.

A: So, do you have any kids?

B: Yes.

While B has technically fully answered A's question, by not elaborating (how many, what sexes and ages) B has not given the amount of information generally expected in response to such a question. As a result, A is likely to get the implicature that B means that she is not comfortable sharing such information with a stranger, without B having to say so directly.

While the next type of implicature, shown in (23c), is perhaps not as common as its opposite, discussed previously in (10c), it is still a possible type. Here the speaker seeks to deceive the hearer into underestimating the difficulty or complexity of the topic of conversation by acting as if the information given, which is actually too little, is sufficient. This may be done to encourage the hearer to attempt something that may otherwise sound too daunting, or to make the hearer feel inadequate when the supposedly easy endeavor turns out to be difficult, or perhaps for other reasons as well. An example of the first reason is given below:

(26)[21] A cookbook is being compiled, and one of the recipes requires the cook to sauté onions on low heat until they are thoroughly browned.

The author writes: This may take 15-20 minutes.

The editor removes this sentence, and writes in a note to the author: Your average reader will be intimidated by 15-20 minutes of stirring onions, but once they start cooking the onions, will continue until they are browned. Remove this time estimate so that it does not keep people from wanting to try this recipe.

Here the cookbook reader ends up simply being told to sauté the onions until they are browned, without any estimate of time. The implicature drawn by the reader is likely to

---

[21] This example is based on a real-life anecdote regarding cookbook editing.

be that the author thinks the time required must not be very significant, or it would have been mentioned. It turns out, however, that this is not true, and the implicature was used deliberately to deceive the reader (and ultimately, to sell more cookbooks by making the recipes look easier).

The last basic type of implicature that may result from giving too little information is that in (23d), which does not involve deceit but is somewhat negative in that it seeks to chastise the hearer and/or cause feelings of remorse or inadequacy. This is the situation type in which the speaker uses underinformativeness to signal to the hearer that the hearer has said something the speaker finds inappropriate.[22] "Inappropriate" here is very broadly defined to include basically anything of which the speaker disapproves, whether because it truly is outside the bounds of decorum, it asks for more information than the speaker wants to give (in this way it overlaps with 23b), it is simply highly uninteresting, or for any other reason. Similar to (23b), this type of implicature involves the speaker sending the hearer the message, "I do not want to talk about this," as in (27).

(27)   Coworkers are chatting during their break.

     A1: So, do you have any pets?

     B talks for several minutes about her current pets and their characteristics, past pets, and pets she hopes to have in the future.

     A2: I have a pet.

Here both B (first) and A (second, in A2) have apparently violated the maxim regarding informativeness: B has given far too much information, and A far too little. The primary violation of concern here is A's provision of too little information, though it is clear that this is in response to B's violation. Leaving aside the possible reasons for B's lengthy answer, it seems that A's very short answer in A2 is intended to send B the message that A thinks her response was inappropriate (because it provided too much information). The implicature, from A's point of view, might be formulated as something like, "I was interested in talking about pets (as evidenced by the fact that I was the one who asked in the first place), but now you've totally overwhelmed me and I don't even want to talk

---

[22] It is worth noting here that flouting of the quality maxim, i.e., sarcasm, is also often used to achieve this aim.

about them anymore," or, "I wanted to talk to you, but then I realized that you apparently have a problem with oversharing, so I no longer want to have this conversation."

As I have shown in this section, both standard and flouting-related implicatures can arise with respect to underinformativeness (or assumed lack thereof). Next, I discuss some properties that hold of all implicatures.

2.2.1.4.3 Eliminating or strengthening an implicature

A hallmark of an implicature is that it can be eliminated without denying the truth of the assertion from which it arose – in Grice's terminology, without denying what was actually "said" – or can be strengthened by asserting what was implicated. This is true of both standard implicatures and those that arise from flouting; while it is unlikely that a speaker would create an implicature through flouting and then choose to eliminate or assert it, the important point is that this could be done. Thus the ability to eliminate or strengthen something serves as a test for whether it is in fact an implicature.

Elimination of a standard quantity implicature can be accomplished by either contradicting or suspending it. Suspension of an upper-bound implicature involves leaving open, but not asserting, "the possibility of something 'stronger' holding;" in the case of scalar implicature, this means that something farther on the right may hold (Horn 1972:45). Thus a suspension of the upper-bound implicature in (16c), i.e., that (16a) and (16b) do not hold, can be accomplished by constructions such as those in (28).

(28)   a. There are some trees in the yard, if not several/many.

   b. There are some trees in the yard, or even several/many.

   c. There are at least some trees in the yard, and possibly even several/many.

(adapted from Horn 1972:40)

All three examples in (28) allow for the possibility that there are several or many trees in the yard, thus suspending the implicature that *some* is the strongest (or right-most on the scale) term that holds in the situation. Possible ways of contradicting an upper-bound implicature are shown in (29).

(29)   a. There are not just some trees in the yard, but several/many as well.

   b. There are some trees in the yard, and moreover / in fact / what's more, there are several/many.

36

In these cases, the implicature that there are not several or many trees is completely eliminated by the assertion that there actually *are* several or many, showing that *some* was clearly not the strongest term that held in the situation.

While it is perhaps intuitively obvious, it is worth noting that the lower bound asserted by quantifiers and scalar predicates is never suspendable, much less contradictable, because it is an assertion rather than an implicature. Thus an utterance such as that in (30a) can only be interpreted as meaning that the speaker has changed his mind about the yard's contents midway through speaking or has realized he is unsure about the contents of the yard, while the remaining utterances in (30) are not acceptable at all.

(30)   a. There are some trees in the yard, or only one.

       b. *There are some trees in the yard, if not only one.

       c. *There are not just some trees in the yard, but only one.

(30a), while arguably acceptable, seems to admit to some degree of unreliability of the speaker. The remaining examples are simply unacceptable by virtue of the fact that they involve near-simultaneous assertion and denial of what was asserted.

Non-scalar standard implicatures can also be suspended or contradicted. For example, as noted previously, a hearer who assumes that the speaker is being neither overinformative nor underinformative will infer that the speaker intends to convey that all of the information he gives is necessary, and that any information not given is unnecessary for the discourse. Adding a statement such as that in (31a) to the end of an utterance can serve to the suspend the former implicature, while (31b) suspends the latter.

(31)   a. …but actually, that might be more than you need to know.

       b. …and there may also be other things you need to know.

(31a) leaves open the possibility that the speaker has in fact given more information than is necessary, while (31b) admits the possibility of too little information having been given. Note also that (31a) and (31b) could be used simultaneously to suspend both standard implicatures. Similarly, the standard implicatures can be eliminated completely by statements such as those in (32).

(32) a. …but you don't really need to know all that.

  b. …and that's some of what you need to know.

(32a) eliminates the implicature that all information given was necessary, while (32b) eliminates the implicature that all necessary information was given. As with the examples in (31), these could be used simultaneously to eliminate both implicatures.

A speaker may also suspend or eliminate an implicature caused by flouting, though as noted above, this is not likely to happen in actual usage. As an example, consider (13), in which A, a violinist, asks B, a luthier, for advice on repairing his bow. B replies with long and detailed instructions and warnings, and by his apparent overinformativeness creates the implicature that the repair process is very difficult and beyond the capabilities of A. However, B could then add something like one of the statements in (33), to either suspend or eliminate this implicature.

(33) a. …but it might not be that hard, and maybe you could do it, so you could give it a try.

  b. …but it won't be that hard; I'm sure you can do it, so you should give it a try.

(33a) serves to suspend the implicature that B's overinformativeness caused, by leaving open the question of whether the task is beyond A's abilities. (33b), on the other hand, eliminates the implicature by asserting its opposite. Implicatures caused by being deliberately apparently underinformative can of course also be suspended or eliminated, with elimination often being accomplished by the speaker providing some reason for his or her underinformativeness other than that which was implicated. Consider (25), in which one stranger (A) asks another (B) whether she has children, and B replies simply with *Yes,* creating the implicature that she does not think it appropriate to share any further information about her children. By following her reply with one of the utterances in (34), B could have suspended or eliminated that implicature.

(34) a. …but since I don't know you, I don't know if you'd want to hear about them.

  b. …and normally I'd tell you all about them, but I have a sore throat today so I'm trying to avoid speaking very much.

(34a) somewhat suspends the implicature by making it clear that B simply is not sure about the appropriateness of sharing further information. (34b) eliminates the implicature

by providing an alternate reason for B's failure to provide further information about her children.

In contrast to suspension and contradiction, assertion of an implicature can be seen as a type of strengthening from assumption to fact; rather than having to assume that something is true, the hearer is told explicitly that this is the case. An implicature can also be strengthened by strongly suggesting, rather than asserting, that what was implicated is true, but this will not be discussed in detail here. Both scalar and non-scalar standard implicatures, as well as implicatures created by flouting, can be strengthened.

Examples of asserting the upper-bound scalar implicature in (16c) are shown below:

(35)   a. There are some trees in the yard, but not several.

         b. There are just/only some trees in the yard.

In (35a), the implicature is explicitly stated as an assertion. The same purpose is served by *just* or *only* in (35b); these words can be thought of as shorthand for something like *and no more* or *and no stronger quantifier holds*. Standard non-scalar quantity implicatures can be strengthened by the addition of assertions such as those in (36).

(36)   a. Remember everything I've told you, because it's all important.

         b. You don't need to know anything else.

(36a) strengthens the implicature that all the given information is necessary by asserting that this is the case. (36b) similarly asserts the previously implicated message that any information not provided is not necessary for the discourse.

Like standard implicatures, implicatures caused by flouting may be strengthened. Returning to the violinist and luthier discussed with reference to suspending or eliminating an implicature, one can see how the implicature created by the luthier could also be strengthened by an utterance such as that in (37).

(37)   …so it's very difficult, and I don't think you should try to do it yourself.

Here the luthier has asserted what was implicated by his apparent overinformativeness. Similarly, an implicature created by deliberate apparent underinformativeness can be strengthened. Returning to the strangers at the bus stop and B's curtness, B could say something like the assertion in (38) to strengthen the implicature she created.

(38)   …but I'm really not comfortable saying anything about them to a stranger.

As in (37), what was previously implicated has been asserted, thus strengthening the implicature. It has been shown, therefore, that scalar and non-scalar standard implicatures, as well as implicatures created by flouting, can be suspended, eliminated, or strengthened.

Grice's (1975) account, together with Horn's (1972) work on scalar implicatures, is arguably able to account for the gap between conventional meaning and what is actually conveyed. Since the publications of Grice (1975) and Horn (1972), there has been substantial further work on the topics of implicature and reducing and refining the maxims of conversation. This is not surprising, given the impact of Grice's work from the beginning and the fact that his work was intended to be only "a general sketch of a potentially operative model, rather than an exhaustive account" (Lakoff 2009:107). Next I discuss three of the more well-known theories built at least in part on Grice's account: Horn's (1984) Q and R principles, Levinson's (1987; 2000) Q-, I-, and M-principles, and Sperber & Wilson's (1986; 1995) relevance theory.

2.2.2 Horn's principles and scalar implicature

Horn (1984) presents a "partial reductionist program" in which Grice's (1975) four implicatures are reduced to principles influenced by the work of Zipf (1949). The two principles are shown in (39).

(39)   Q and R principles

a. The Q principle

Make your contribution sufficient.

Say as much as you can (given R).

b. The R principle

Make your contribution necessary.

Say no more than you must (given Q).   (Horn 1984:13)

Unlike Grice's (1975) maxims (which were nonetheless acknowledged to interact), Horn's principles explicitly reference each other, making it obvious that they interact. The Q principle is intended to capture the first part of the quantity maxim (an injunction to provide enough information) and the first two parts of the manner maxim (injunctions against obscurity and ambiguity) (Horn 2004). The R principle covers the rest of the

maxims, with the exception of quality: relation, the second part of the quantity maxim (regarding not saying too much), and the last two parts of the manner maxim (regarding briefness and orderliness). The quality maxim, while not formulated into a named principle, is retained with its special status as the most important maxim (mentioned in footnote 17). Horn (1984) argues that the Q and R principles, along with the quality maxim, comprise a reduced theoretical apparatus that covers at least as wide a range of pragmatic and linguistic phenomena as Grice's (1975) work.

Like Grice, Horn (1972; 1984) fails to explicitly address several questions about his view of informativeness. For example, it is again unclear whether only propositions may be informative, or other items may be as well. Horn's (1972) characterization of implicature in terms of entailment suggests a requirement of propositionality, since only a proposition can entail or be entailed, but the issue is not addressed in Horn's (1984) introduction to the Q and R principles. Also unaddressed in that work is the question of whether an addressee's previous knowledge is considered with relation to informativeness; as with Grice's (1975) quantity maxim, Horn's (1984) Q principle certainly seems to leave open the possibility that the addressee's knowledge is taken into account. Further, Horn's (1972) account of presupposition suggests that Horn does in fact place some importance in general on an addressee's cognitive state prior to an utterance. Other than the fact that his work is based on that of Grice, there is no indication that Horn views truth as a necessary characteristic of information (though it is retained as a principle of conversation), but there is likewise no indication that he does not. Regarding the independence of informativeness and other utterance attributes, Horn's (1984) account resembles that of Grice (1975) in that utterance attributes such as clarity and conciseness are not considered directly as a part of informativeness but are addressed by parts of his Q and R principles. Thus in his treatment of informativeness, Horn (1972; 1984) differs little from Grice (1975).

2.2.3 Levinson's Q-, I-, and M- principles

Levinson (1987) examines the work of Zipf, Horn, Grice, and others in some depth, and builds on the work of Atlas and Levinson (1981). He focuses specifically on generalized conversational implicatures, and proposes the Q- and I-principles, each

41

containing a "speaker's maxim," outlining what the speaker should and/or should not do, and a "recipient's corollary," explaining what the principle "licenses the addressee to think" (Levinson 1987:67). He does not explicate the M-principle, but acknowledges its probable necessity and distinguishes between Q-implicatures and Q/M-implicatures. The M-principle is formally proposed in Levinson (2000). Slightly condensed versions of all three principles are shown in (40).

(40) Q-, I-, and M-principles

   a. The Q-principle

   Speaker's maxim: Make your contribution as informative as is required for the current purposes of the exchange. Specifically: don't provide a statement that is informationally weaker than your knowledge of the world allows, unless providing a stronger statement would contravene the I-principle.

   Recipient's corollary: Take it that the speaker made the strongest statement consistent with what he knows, and therefore that:

   > (i) If the speaker asserted A(W), and <S, W> form a Horn scale, then one can infer that the speaker knows that the stronger statement A(S) would be false.

   > (ii) If the speaker asserted A(W) and A(W) fails to entail an embedded proposition q, which a stronger statement A(S) would entail, and S and W are 'about' the same semantic relations, then one can infer that the speaker doesn't know that q obtains.

   b. The I-principle

   Speaker's maxim: The maxim of minimization

   Say as little as necessary to achieve your communicational ends, bearing Q in mind.

   Recipient's corollary: Enrichment rule

   Amplify the informational content of the speaker's utterance, by finding a more specific interpretation, up to what you judge to be the speaker's m-intended[23] point. Specifically:

---

[23] "M-intended" is used here as defined in footnote 16.

(i) Assume that stereotypical relations obtain between referents or events, unless this is inconsistent with what is taken for granted or the speaker has broken the maxim of minimization.

(ii) Assume the existence or actuality of what a sentence is 'about,' if that is consistent with what is taken for granted.

(iii) Avoid referential interpretations that multiply entities in the domain of reference.

c. The M-principle

Speaker's maxim: Indicate an abnormal, nonstereotypical situation by using marked expressions that contrast with those you would use to describe the corresponding normal, stereotypical situation.

Recipient's corollary: What is said in an abnormal way indicates an abnormal situation, or marked messages indicate marked situations, specifically:

(i) Where S has said something containing a marked expression, and there is an unmarked alternate expression U with the same denotation D which the speaker might have employed instead, then where U would have implicated the stereotypical subset d of D, the marked expression implicates the complement of d.

(Levinson 1987:67-8; 2000:136-7)

The Q-principle covers the first part of Grice's (1975) quantity maxim (an injunction to provide enough information), while the I-principle captures the second part (regarding not saying too much). As with Horn's Q and R principles, these two refer to each other. The M-principle, according to Levinson (2000), covers two or perhaps three parts of the manner maxim; he questions whether orderliness should have maxim status, and claims that avoiding ambiguity is partially covered by the Q-principle. This leaves, of course, the maxims of relation and quality; relation is set aside as applicable to particularized conversational implicatures only, while quality "plays only a background role in the generation of" generalized conversational implicatures, and is thus excluded from the three principles (Levinson 2000:74). Levinson (2000) also notes that his principles can be compared to Horn's (1984) Q and R principles, with Levinson's Q- and M-principles

43

together corresponding to Horn's Q principle, and Levinson's I-principle being roughly equivalent to Horn's R principle. Equivalencies notwithstanding, Levinson (1987; 2000) motivates his account through examinations of and establishment of connections between Gricean inference, conversation analysis, syntax, and the semantics-pragmatics distinction.

In describing the properties of information and informativeness according to his account, Levinson (2000) is not much more explicit than Horn (1972; 1984) or Grice (1975). Like Horn (1972), Levinson (2000) characterizes informativeness in terms of entailment, indicating that propositionality is at least implicitly required for informativeness. As with Grice's (1975) quantity maxim and Horn's (1984) Q principle, Levinson's (1987) Q- and I-principles seem to show that an addressee's prior knowledge should be taken into consideration. Most specifically, in order to "say as little as necessary to achieve your communicational ends," a speaker would have to know what prior information the addressee has in order to avoid saying something that was not strictly necessary. Levinson offers no clues about his view of whether information must, by definition, be true. Finally, as with the accounts of Grice (1975) and Horn (1984), it is clear from Levinson's (2000) M-principle that other utterance attributes are considered important but are not part of informativeness. In this way, Levinson's view is also similar to relevance theory, the next (and last) theory to be discussed.

2.2.4 Relevance theory

Sperber and Wilson's (1986; 1995) relevance theory can be seen as the most extreme example of the reductionist trend shown by Horn and Levinson: while Horn (1984) reduces Grice's four maxims to two principles, and Levinson (1987; 2000) to three, Sperber and Wilson (1986) initially propose just one principle: that of relevance (they later revise this to two principles). In relevance theory (RT), speaker intention and audience recognition of that intention are important. RT is based in part on the premise that neither the code model (as in information theory) nor the inferential model (as in Gricean implicature) is sufficient to explain human verbal communication, but the two may be combined to form a satisfactory account.

Another premise of RT is that humans manage their cognitive processes in such a

44

way as to maximize effects, and particularly positive cognitive effects, while minimizing effort; this basic idea is common to those building on the work of Zipf (1949) (Sperber & Wilson 1995). Positive cognitive effects are defined as the modification of a cognitive context, or cognitive environment, in a way that "improves" it, by interaction between old and new information, strengthening of assumptions, or contradictions that result in the removal of assumptions (Sperber & Wilson 1995). One's cognitive environment, or the set of facts that are manifest to him, "can be modified by adding a single piece of new information, but it can equally well be modified by a diffuse increase in the saliency or plausibility of a whole range of assumptions, yielding what will be subjectively experienced as an *impression*" (Sperber & Wilson 1995; Wilson & Sperber 2012:87). In attempting to create positive cognitive effects in an addressee, therefore, a speaker must take into account what information the addressee may already have, in order to know what new information may improve the addressee's cognitive context. Wilson and Sperber argue that a speaker "can often predict… which background information from [the addressee's] memory is likely to be retrieved and used in processing [the speaker's] stimulus" (2012:272).

Cognitive effort is seen to be proportional to the amount of conceptual material in the stimulus to be processed; in this way RT takes into account attributes of an utterance such as clarity and conciseness, and they factor into relevance but not informativeness (Sperber & Wilson 1995). Processing implications also requires cognitive effort, but Sperber and Wilson argue that this effort is immaterial to the calculation of relevance because it "is always in proportion to the effects it implements" (1995:126).

The minimization of effort and maximization of effect proposed in RT is done by providing the most attention to "what seems... to be the most relevant information available," where 'relevant' is not yet defined but is presumably used in its general, non-technical sense (Sperber & Wilson 1995:vii). Because human verbal communication places demands on attention, "to communicate is to imply that the information communicated is relevant" (Sperber & Wilson 1995:vii). From this premise is extracted the original principle of relevance, shown in (41).

(41)   Principle of relevance

> Every act of ostensive communication communicates a presumption of its own optimal relevance. (Sperber & Wilson 1995:158)

In their (1995) postface, Sperber and Wilson refer to the principle in (43) as "the second principle of relevance," and designate the principle in (42) as the first.

(42)   The first principle of relevance

Human cognition tends to be geared to the maximisation of relevance. Relevance itself is defined as the degree to which an assumption in a context has contextual effects that are large combined with the degree to which it requires a small amount of processing effort; "…an utterance is *relevant* when the hearer, given his cognitive dispositions and the context, is likely to derive some genuine knowledge from it" (Sperber & Wilson 1995:125, Wilson & Sperber 2012:60). Sperber and Wilson use 'assumption' to refer to any thought, whether represented as true or not: they contrast their view of information with that of some others, noting that RT does not limit information to that which is true; instead, they "use the terms [information and inform] more broadly, treating as information not only facts, but also dubious and false assumptions presented as factual" (1995:2). In later work, however, they appear to revise this stance; in their 1995 postface, they note that a theory that takes no account of "objective truth or falsity" is "incomplete," but claim that it is "good enough… for most of our purposes" (Sperber & Wilson 1995:263); in 2012, they state their agreement with Grice's statement that "false information is not an inferior kind of information; it just is not information" (1989:371), and follow with the conclusion that something that is relevant is "therefore true" (Wilson & Sperber 2012:83).

In terms of measurement, relevance is treated comparatively rather than quantitatively; Sperber and Wilson claim that relevance could be defined quantitatively, but state, "since we are interested in relevance as a psychological property, we have no reason to aim for a quantitative definition of relevance" (1995:132). They define an "informative intention" as the intention of a speaker "to make manifest or more manifest to the audience a set of assumptions I" (1995:58). Something is considered manifest relative to a particular individual and time; an assumption is "manifest to an individual… if he is capable at that time of representing it mentally and accepting its representation as

true or probably true" (Sperber & Wilson 1995: 39). Increasing manifestness is one way of achieving the positive cognitive effects associated with relevance. With the definitions given here for positive cognitive effects and manifestness, it should be clear that, from a relevance theoretic standpoint, a recipient's knowledge or beliefs about the reliability of a source are important: a recipient who is less confident in the reliability of a message, regardless of the message's actual reliability, will likely have her cognitive environment less affected by that message because she has less confidence in the reliability of the information produced. The net effect is that, ceteris paribus, under RT a message will be less informative when received by someone with less confidence in the source from which the message originated.

Given that an informative intention involves an intention to increase manifestness, the informativeness of a message is a function of the degree to which it increases the manifestness of some set of assumptions. Note that this means that if one is already absolutely certain about a set, then having it further confirmed achieves no positive cognitive effects and is completely uninformative, much like a message containing only information already possessed by the recipient is considered uninformative in IT and, it would seem, in most other theories (Sperber & Wilson 1995).

In Sperber and Wilson's (1995) discussion of informativeness, there seems to be an implicit assumption that the messages that may be informative are propositional. However, they note that, although pragmatists have generally confined their treatments of "meaning" to propositional communication, "much of what is communicated does not fit the propositional mould" (Sperber & Wilson 1995:57). Regarding non-propositional items of a variety of types, including non-linguistic ones, Sperber and Wilson "see it as a major challenge for any account of human communication to give a precise description and explanation of its vaguer effects" and assert that "once this step is taken… the framework we propose, unlike the others we have discussed, can rise to this challenge" (1995:57-58). However, as noted previously, they did later express the view that information must be true; it seems, therefore, that either relevance is not limited to propositions, but informativeness is (a rather strange claim, I think), or Wilson and Sperber (2012) revised their opinions on propositionality. Thus relevance theory

originated based on a propositional model, and may or may not remain confined to propositions. Thus far, it has typically been used in the examination and assessment of propositional items, so its non-propositional capabilities, if any, remain largely unexplored.

Their two principles of relevance, Sperber and Wilson (1995) argue, are sufficient to provide an account of communication that is more robust than either the Gricean inferential model or an information-theoretic code model alone, and that can explain implicatures based on all of Grice's (1975) maxims. However, some disagree, claiming that an adequate theory requires reference to more than just one of Grice's original maxims and that Sperber and Wilson have taken reductionism too far (e.g., Levinson 1987; Gundel & Mulkern 1998; Lakoff 2009).

2.3 Conclusion

In this chapter I have intended to provide a reasonably complete, if somewhat brief, account of how information and informativeness have been used in some of the major theories both within and outside of linguistics. Within linguistics, I have focused on conversational implicature in order to illustrate one of the main areas of linguistics in which the concept of informativeness is used. First proposed by Grice (1975) in terms of the cooperative principle and conversational maxims, the notion of implicature has seen much work by later scholars such as Horn (1972; 1984; 2004), Levinson (1987; 2000), and Sperber & Wilson (1995; 2012). Implicatures arise whether maxims (or, in later work, principles) are obeyed or violated. Further, they arise whether violation is intentional or not, though the character of the implicature varies based on the perceived intentionality of the violation. I have characterized implicatures arising from obeyed maxims as "standard" implicatures, following Levinson (1983), and have used the conventional term "flouting" to refer to the process by which implicatures are created through intentional violation of a maxim. Both types of implicatures can be suspended, eliminated, or strengthened, and this serves as a test to distinguish what is implicated from what is asserted.

In the theories associated with implicature, informativeness has most often been characterized in terms of entailment: if x entails y, but y does not entail x, then x is more

48

informative than y. It is thus a relative measure rather than an absolute, quantitative one; this is fairly standard among measurements of informativeness in linguistics, as is the appeal to entailment. However, some who have written about implicature, such as Grice (1975), have omitted any explicit definition of informativeness, while others, such as Hirschberg (1985), have proposed types of quantity implicature that do not involve entailment. Overall, then, the topic of implicature is somewhat representative of the field of linguistics as a whole when it comes to informativeness: explicit definitions of informativeness are often left out; when it is defined, informativeness is frequently characterized in terms of entailment; and some dispute this definition or use very different ones without comment on the entailment characterization. In the next chapter, I give an account of a study I did that led to the need for an account of informativeness, and then describe the characterization I formulated.

Chapter 3. Informativeness: motivation and characterization

3.0 Introduction

As mentioned in chapter 1, the initial motivation for the idea of a category membership-related measure of informativeness was the distribution of English adjectival past participles (APPs) in attributive position. In this chapter I recount the APP study in order to partially illustrate the motivation for the present work. The study outlined here also serves as an example of another area in which the concept of informativeness is useful; however, here it is used not with the goal of explaining the mismatch between what is said and what is communicated, but with the goal of explaining why some expressions are more used than others.

3.1 APP study

In van Duym (2010) I examined the distribution of two particular forms of the English adjectival past participle: the regular (non-negated) form, e.g. *flown*, and the form negated with the prefix *un-*, e.g. *unflown*. These were compared in prenominal attributive position within noun phrases, e.g. *the unflown plane*.

Many APPs that are able to function attributively can take the negative prefix *un-*. This prefix can have either of two distinct meanings, the first indicating a reversal, and the second indicating the absence of a state or property. The former meaning of *un-* is relatively rare in the data used here, and almost always coexists ambiguously with the latter meaning.[24] Because instances of the prefix indicating a reversal are so rare in the data and do not seem to be relevant, the distinction between these two senses of *un-* will not be further discussed.

Examination of relative frequencies showed that there were at least two, and perhaps three, distinct patterns of distribution between non-negated and negated forms: some APPs occur primarily without negation, a few appear equally often with and without the negating *un-*, and a final group is found most often with *un-*.[25] To explain why these

---

[24] For example, *unrolled* in *The unrolled rug took up more room than they had expected* can also be interpreted as meaning that the rug had never been rolled , or that the rug had been rolled, and then the process was reversed.

[25] For other studies that address adjectival past participles requiring some kind of additional modifying element to be acceptable, see Ackerman & Goldberg (1996), Grimshaw & Vikner (1993), Wasow (1977) and Lakoff (1970).

distribution patterns exist, I examined factors related to semantic fields, situation aspect, incremental theme, historical origins, and informativeness. I concluded that the most significant factor was informativeness, and while none of the examined factors was solely able to explain the distribution patterns of negated and non-negated APPs, all at least potentially contribute to a multifaceted explanation that likely also involves aspects that remained unidentified.

3.1.1 Methodology and data

Having established that APPs in general occur in noun phrases both with and without the *un-* prefix, I compared the distribution of the prefixed and unprefixed forms for a wide variety of APPs. The Corpus of Contemporary American English (COCA) was used as the data source and was searched for each form within a phrase containing *a*, *an,* or *the* followed by the APP and then any noun.[26] For each query, the total number of occurrences for the first 1000 tokens was recorded; in most cases this was equivalent to the total occurrences overall, as very few searches yielded more than 1000 tokens.[27] For each APP, the proportion *r* of total non-negated to negated forms was calculated by dividing the total number of occurrences of non-negated forms *p* by that of negated forms *n*, i.e., $r = p/n$.[28] Thus APPs that occurred more often in negated than non-negated form had $r < 1$, and APPs that occurred more often in non-negated form had $r > 1$. An *r* value of 1 indicated that an APP occurred equally often in negated and non-negated forms.

204 APPs were tested, each in both non-negated and negated form, for a total of 408 search queries. Initially APPs were chosen somewhat randomly as they came to mind or were encountered in various contexts not directly related to this research. However, once a pattern of distribution became apparent, I began to seek out the more elusive APPs that

---

[26] The requirement of one of these three determiners before each APP served several purposes: first, it somewhat restricted an extremely large data pool (which might, of course, be viewed as good or bad depending on one's perspective); second, it helped to ensure that the phrases retrieved were, in fact, noun phrases; finally, it reduced (though did not eliminate) the number of retrieved phrases containing plural nouns, which can be problematic for situation aspect (discussed in 3.1.3).

[27] As used here, a *token* means a specific phrase in which an APP was found; an *occurrence* is a specific instance of a token being found in the corpus. So, for example, for the APP *unbroken* there are 128 tokens, including *an unbroken string*, *the unbroken chain*, and *the unbroken record* (plus 125 others). For the token *an unbroken string*, there are 6 occurrences, meaning that the phrase is found 6 times in COCA.

[28] All *r* values were rounded to one decimal place.

occur predominantly with the *un-* prefix, selecting verbs to test that I thought (generally through my own native speaker intuition) might fulfill this criterion. The number of APPs with $r < 1$, indicating that they occurred more often in negated than non-negated form, was 29, or 14% of those tested. Henceforth these will be referred to as 'negative APPs'. This data can be seen in the first column of Table (1) below, and confirms my initial intuition that such forms are a minority, especially considering that this was not a random sampling but one in which these types of APPs were actively sought. Another 26 APPs, in the second data column of Table (1), had r in the 1.0-2.2 range, indicating that their non-negated forms did not strongly predominate over the negated ones. It was unclear whether these were best grouped with the lower or higher r-value APPs, or whether they constituted a distinct distributional group on their own, and this has not been further explored. 2.2 was chosen as the upper bound for this group because it reflected a natural break in the r values of the data. Finally, the vast majority of APPs had r values greater than or equal to 2.6, indicating strong predominance of the non-negated form over the negated form; or r values that could not be calculated due to an n of 0, indicating that the negated form was not found at all in the relevant context in the corpus. Henceforth these APPs with $r \geq 2.6$ or undefined will be referred to as 'positive APPs.' The data is summarized in Table 1.[29]

	$r = 0.0\text{-}0.9$	$r = 1.0\text{-}2.2$	$r \geq 2.6$ or $n = 0$ (r undefined)
# of APPs	29	26	149
% of tested APPs	14%	13%	73%

Table 1: APP distributional data

Given this data, an obvious question arose: what characteristics of this minority, the negative APPs, caused them to occur predominantly with the negative prefix *un-*, as shown in the first column of data? Several factors, beginning below with semantic fields, were examined in an attempt to answer this question.

3.1.2 Semantic fields

The first factor examined was semantic, as it became apparent very early on in the

---

[29] Percentages here, and throughout this paper, are rounded to the nearest whole percentage point.

research that many of the negative APPs seemed to share a small number of semantic fields or conceptual categories. Specifically, words having to do with mental processes, such as *unknown*, *undoubted*, and *unforeseen*, appeared to be overrepresented in the negative APP category. Other categories represented by several negative APPs included speech-related activities (*unsung*, *unspoken*) and acts involving physical movement (*unzipped*, *unshaven*). However, members of these three fields were also found among the positive APPs (though in lesser proportions). The fact that these categories were represented in both sets of APPs strongly suggested that these semantic distinctions alone were not sufficient to explain the differences in distribution. However, it is also true that the identified semantic fields were represented to a much greater degree among the negative APPs than the positive ones. This indicated that the identified categories may play a small part in explaining the distributional pattern, if indirectly.

3.1.3 Situation aspect

The next potentially explanatory factor examined was situation aspect; I based my analysis of this primarily on Smith (1997). Though it showed some promise, the viewing of this data through a situation aspect lens was problematic. The main reason for this was that situation aspect, as discussed by Verkuyl (1972), Smith (1997) and others, has to do most prominently with a verb; indeed it has been treated in the past as relating solely to the verb (Smith 1997). As noted earlier, the past participles examined here were functioning not as verbs but as adjectives. Thus aspect could only be discussed with reference to the verbs from which these APPs were derived, rather than the APPs themselves (Grimshaw & Vikner 1993). In addition, many linguists now accept that aspect cannot be determined based solely on a verb, but rather requires what Smith refers to as a "verb constellation," which includes not only the verb itself but its arguments (Smith 1997:5,17; Verkuyl 1972). As adjectives, APPs did not have the types of arguments found in a verb constellation, nor was there always one clear (generic) set of arguments that the verb from which an APP was derived would have. These factors made talking about situation aspect in relation to APPs messy at best, but they did not render the concept useless.

In order to assess the type of situation aspect that might be associated with each APP,

I considered the verb from which it was derived together with the arguments suggested by the noun phrases in which the APP occurred; generally, this meant a subject and a singular object. I avoided using plural objects because they can cause atelicity where it would not otherwise be found, due to an unspecified number in the object preventing a natural endpoint from arising (cf. *eat an apple* vs. *eat apples*) (Smith 1997:55). Using this method, I attempted to assess whether the verb that each APP was derived from was static – comprising a single, undifferentiated period – or dynamic with successive stages; instantaneous (conceptually, at least), or durative; and telic – having a change of state which constitutes the outcome or goal of the event – or atelic (Smith 1997:19).

Very few of the positive or negative APPs were derived from static verbs, with most coming from dynamic verbs instead: only four of the negative APPs and eight of the positive APPs had associated verbs that were conclusively identified as static. A large portion (48%) of the negative APPs came from exclusively durative verbs, as opposed to 28% from exclusively instantaneous verbs.[30] In contrast, a smaller portion of the positive APPs (40%) were derived from verbs of exclusively durative classification, and this was matched by 40% derived from verbs of exclusively instantaneous classification. When looking only at verbs whose $n=0$ (a group included in the positive APPs), the percentage derived from durative verbs is an even lower 26%, compared to a much higher instantaneous verb percentage of 64%.

The telicity of some APPs' associated verbs could not be definitively determined; among the APPs classified, all groups had more telic than atelic associated verbs. Among positive APPs, 47% were associated with telic verbs and 27% with atelic verbs; among negative APPs, 62% were associated with telic verbs and 31% with atelic verbs. Due to the fact that significantly more positive than negative APPs were unclassifiable (26% versus 7%), it is difficult to say whether my expectation that negative APPs' associated verbs would have a higher rate of telicity was in fact confirmed. A more robust method of determining situation aspect characteristics of verbs from which APPs were derived would likely have been needed in order to assess whether telicity played an explanatory

---

30 Many APPs, both positive and negative, appeared in multiple senses (whose durativity varied) in the data and as such could not be classified as exclusively durative or instantaneous.

role in the distribution of APPs. This was not pursued at the time of the study, and remains a potential direction for further research on APPs.

In summary, positive APPs were most likely to be derived from verbs that were dynamic and telic; assuming that these characteristics occurred together, these were verbs that, by Smith's (1997) classification, would be considered either accomplishments or achievements. In slight contrast, negative APPs were most likely to be derived from verbs that were dynamic, durative, and telic; Smith (1997) would classify a verb having all these characteristics as an accomplishment. In fact, achievement proved to be the most common classification of APPs' associated verbs overall.

Taken as a whole, this data suggested that situation aspect alone was not highly explanatory of the distribution of positive and negative APPs. What this data did show was that APPs that could be used in the position studied here, both positive and negative, were often derived from verbs of the accomplishment or achievement type. This was significant largely because the accomplishment type has previously been identified as contributing to modification requirements such as that satisfied by the use of *un*- (Grimshaw & Vikner 1993). According to Grimshaw and Vikner, all accomplishment verbs have a two-part event structure consisting of a process and a resultant state (1993:144-145). An "obligatory adjunct," a term which they use rather broadly to also include negating prefixes, serves to "identify" one of the sub-events that would otherwise go unidentified (1993:144). They show that this is required whenever the accomplishment verb is one involving creation, or what they call "constructive accomplishments," because in other types of accomplishments, there is not an unidentified subevent (Grimshaw & Vikner 1993:146-147).[31] Grimshaw and Vikner would predict that I should have found no constructive verbs in the positive APP category, given that the noun phrase context in which I searched the corpus eliminated obligatory adjuncts of other types. However, this was not what I found; this strongly suggested that something was going on beyond what Grimshaw and Vikner identified, and motivated me to continue looking at other factors.

---

[31] Constructive accomplishments are defined as ones in which "the Theme did not exist in its present form before the event occurred," though this is intended to exclude situations in which the new "form" of the theme is the result of a destructive process (Grimshaw & Vikner 1993:146).

3.1.4 Incremental theme

The notion of incremental theme, first proposed by Krifka (1987, 1989), Dowty (1987) and Hinrichs (1985), was the next factor examined. An incremental theme is a theme whose parts are mapped onto parts of an event by a telic predicate. This allows for a further distinction in the situation aspect realm than those made by the criteria discussed with reference to Smith (1997) above (Dowty 1991:567). Both positive and negative APPs had associated verbs that occurred with incremental themes, but both also had verbs with non-incremental themes. However, there seemed to be a possible trend that was consistent with my expectations: the verbs from which positive APPs were derived had a slightly lower proportion of incremental themes among them than did the verbs from which negative APPs were derived. More precise and comprehensive identification of incremental themes was still needed to say with certainty whether they play a role in the distribution of positive and negative APPs; this was done later and is further discussed in 3.2.4.

3.1.5 Word origins

Another element of a rather different type examined in this study was historical provenance. Because the negative prefix being used, *un-*, is of Germanic origin, it was reasonable to consider whether the negative APPs, which take the prefix much more commonly than the positive ones, also more commonly share a historical background with the prefix. The reasoning behind this was, in part at least, that sharing a common origin with the prefix meant a word had existed in the same language as the prefix for a longer time than a word with unrelated origins. This could correlate with the prefix being more productive with that word, and with other negative prefixes being therefore less likely to be affixed to the word. It was not clear that historical factors could explain the lack of positive instances for some forms; origin was offered solely as a possible reason for the abundance or scarcity of *un-* prefixed negative forms relative to the total number of negative forms produced in any way.

Historical information was sought in *The Oxford English Dictionary* (1989) for the verbs from which all studied APPs were derived, and was available for the vast majority of them. My prediction that a higher percentage of the verbs from which negative APPs

were derived would be Germanic than Latin/French, and that the reverse would be true for verbs from which positive APPs were derived, was initially strongly supported by a smaller data set. With the final data set, however, this was supported only very weakly. Though by no means conclusive or sufficient to solely explain the patterns of distribution, historical information suggested that word origin could be yet another factor that played a meaningful role in the use of *un-* on APPs.

3.1.6 Informativeness

The last factor examined was borne of the observation that many of the forms found infrequently or not at all seemed, to some degree, semantically or pragmatically anomalous. Under usual circumstances,[32] why would one refer to a man as *the identified man*, rather than just *the man*, when being identified is part of the default or assumed state of a man? Why would one say *the unbent wire* when the equivalent phrase *the straight wire* is available? I grouped the answers to these and similar questions under the umbrella of informativeness.

It is generally recognized by Griceans and neo-Griceans alike that, under most definitions of informativeness, speakers typically prefer the more informative of two otherwise similar utterances of comparable length. Likewise, given two utterances of similar informativeness, the more concise one is usually preferred. These preferences can be seen in Atlas and Levinson's statement that, "given that there is available an expression of roughly equal length that is…more informative, the failure to employ [it] conveys that the speaker is not in a position to employ it" (1981:38). They are also reflected in Grice's (1975:45-6) maxims of manner and quantity (specifically: "be brief;" "make your contribution as informative as is required"), Horn's (1984:13) Q and R principles (specifically: "say as much as you can; say no more than you must"), and Levinson's (1987:67-8) Q- and I-principles (specifically: "don't provide a statement that is informationally weaker than your knowledge of the world allows;" "say as little as necessary to achieve your communicational ends"). Given language users' preferences for conciseness and informativeness, one should expect that a phrase of the type studied

---

[32] Henceforth, I use the phrases 'usual circumstances' and 'neutral circumstances' to refer to situations not greatly affected by contextual or confounding factors aside from any currently being discussed.

here was rarely or never found if it was no more informative than the corresponding simple [Determiner + Noun] phrase without the APP, since then the phrase without the APP should occur instead. So if an APP was primarily or always found in negated form, it was predicted that the negated form typically contributed to the informativeness of the phrase in which it occurred, while the non-negated one did not (or did so less). Likewise, if an APP was primarily or always found in non-negated form, it was expected that that form typically contributed to informativeness, while the negated form did not (or did so less).

Informativeness (as well as information) has been defined in various ways, as discussed in chapter 2. Though it necessitates explicit defining of information and informativeness by researchers, can make consistencies difficult to identify, and can discourage cross-theoretical comparison and knowledge transfer, the flexibility with which the concepts are used is convenient in that it allows each researcher some freedom to customize characterizations that best work with the research at hand. Thus I now turn to explicitly showing how I used 'information' and, in particular, 'informativeness' in this context.

Recall that a recurring theme in characterizations of informativeness is the elimination of possibilities. Continuing this theme in the context of theories of categorization, Giora states, "to say that a category member is informative in a given set is to say that it has more features (information) than necessary for category inclusion" (1988:550). E.g., for distinguishing members within the set or category of apples, the phrase *an apple smaller than a house* is not very informative, because it describes a characteristic of all apples[33] that might thus be assumed of anything that is included in the apple category.[34] In contrast, *a green apple* is much more informative because it describes a feature beyond what is necessary to be considered an apple; certainly not all

---

[33] Note that category members are limited to actual referents of the category name, and exclude representations thereof. E.g., there are water towers whose tanks are made to look like apples, and some of these are larger than houses, but they do not fall within the apple category. For more on this, see chapter 4, footnote 2.

[34] Should someone happen to develop an actual apple larger than a house, this assumption might be weakened slightly but would remain; the fact that we may recognize something otherwise apple-like but larger than a house as an apple suggests that smallness is not, in fact, a strict requirement for inclusion in the category of apples.

members of the apple category are green, so knowing that an item is an apple does not constitute knowing that it is green. Thus if one knows that an item belongs to the category of apples, one also knows that it is smaller than a house, but not whether it is green. Giora's (1988) characterization is not unlike Sperber and Wilson's (1995) treatment of positive cognitive effects, as discussed in 2.2.4 – Giora's informativeness involves providing information about a category member that one does not already have by virtue of knowing its category membership, while positive cognitive effects can be achieved by providing information that one does not already have but that is related in some way to information that one does already have. Thus informativeness within a category can be viewed as a specific type of situation in which positive cognitive effects are generated. For the APP study I developed a characterization of informativeness influenced by the work of Giora (1988) and Sperber & Wilson (1995): a phrase or utterance is informative to the extent that it increases the cognitive availability of information not already fully available via category membership or context, and thereby reduces the number and/or likelihood of possibilities for the event, state, or entity to which it refers.

Of course, some elaboration on this characterization was necessary. I initially sidestepped the issue of propositionality via a proposed conversion of noun phrases into propositions, so for the purposes of the APP study I, like Dretske (1981) and others, confined informativeness to propositional items. Unlike Dretske (1981), Grice (1989), and Wilson & Sperber (2012), I did not require that information be true; merely being "presented as factual" and capable of being represented as "true or probably true," as Sperber and Wilson included in their early definitions of information and manifestness, was sufficient (1995:2,39).

In characterizing exactly which propositions were considered informative, I relied in part on the familiar theme of reduction of uncertainty. I also employed the notion of positive cognitive effects; recall from chapter 2 that these are defined as effects on the cognitive environment resulting from interaction between old and new information, the strengthening of assumptions, or the processing of contradictions that result in the removal of assumptions (Sperber & Wilson 1995). Note that it may be possible to have a

positive cognitive effect without having a reduction of uncertainty; if a new piece of information causes a contradiction and the result is that both the new piece of information and the previous assumption with which it contradicted are cancelled (or at least less trusted, and thus made less manifest, in the language of Sperber & Wilson 1995), then a positive cognitive effect has occurred but there has been no reduction (and possibly even an increase) in uncertainty. It is for this reason that I included both concepts in my characterization of informativeness, but with an emphasis on positive cognitive effects, so that in a situation where positive cognitive effects occurred but there was an increase in uncertainty, informativeness was still achieved. As suggested by the reference above to manifestness, I also included this as a way that a proposition may be informative: by increasing the manifestness of some set of assumptions.

Finally, as most theories do, I excluded from the informativeness assessment any information already possessed by the message recipient; this accounts for the "not already present via category membership or context" portion of the characterization above. I did not heavily restrict context, as is done in information theory. I treated context in a way similar to Sperber and Wilson (1995), so that context comprised the entire cognitive environment, and thus included an individual's knowledge deriving from cognitive, physical, cultural, and discourse-related sources.[35] This was done with the understanding that language users have the capacity to determine what aspects of context are relevant in a given situation. That is, a language user need not consciously consider everything she knows, can sense, and has experienced every time she creates or interprets an utterance. With this definition, prior information from context encompasses all of what was referred to as from "category membership or context" above, and all prior information the recipient might possess can be said to come from context, so that "not already present via category membership or context" can be reduced to "not already present via context". I chose not to word my account of informativeness this way initially in order to highlight the information that comes from category membership, as this was a major focus of the original study; it has since become an even more prominent aspect of my work, so the

---

[35] It is important to note that some aspects one's physical environment, for example, are not part of one's cognitive environment, because one is simply not aware of them; the reliance on cognitive environment is meant to exclude consideration of anything of which one is completely unaware.

foregrounding of it remains useful. I conformed with relevance theory in suggesting that informativeness be measured comparatively rather than quantitatively. Unlike Sperber and Wilson (1995), however, I do think that a quantitative measure could be quite useful under certain circumstances, and efforts to develop such a method appropriate to one's research are not wasted.

Assessing exactly what prior information a language user has is important to the calculation of informativeness as defined here. Because a unique facet of my characterization of informativeness was its explicit reference to information provided by category membership, as part of the APP study I proposed the beginnings of a method for assessing that information, outlined below. It is the further development of that method in general, rather than with reference to APPs specifically, that is the focus of most of the rest of this dissertation.

For the assessment of what information can be said to be provided by category membership, I first considered that information about any characteristic known to be necessary for category inclusion would be fully cognitively available as a result of knowing an item's category membership; as Smith and Medin note, categories "allow us to go beyond the information given" and "infer…attributes" (1981:1). In other words, if I know that an item is a member of the apple category, and I also know that, in order to be a member of that category, an item must be a fruit, then the information that the item in question is a fruit is fully cognitively available to me. In accounting for other information provided by category membership, I utilized Pustejovsky's notion of qualia structure: "the structured representation which gives the relational force of a lexical item" and constitutes "modes of explanation for a word" that "permit a much richer description of meaning" than a purely decompositional or relational view would provide (1998:76). The qualia that Pustejovsky proposes for a nominal are shown below:

(1) Pustejovskian qualia

  a. Constitutive: the relation between an object and its constituent parts

   Material; weight; parts and component elements

  b. Formal: that which distinguishes it within a larger domain

   Orientation; magnitude; shape; dimensionality; color; position

c. Telic: its purpose and function

    Purpose that an agent has in performing an act; built-in function or aim that specifies certain activities forming an act

d. Agentive: factors involved in its origin or bringing it about

    Creator; artifact; natural kind; causal chain

(Pustejovsky 1991:426-7; 1998:76)

Pustejovsky notes that "not all lexical items carry a value for each qualia role," and a lexical item may have more than one quale of any given type (1998:76; 1991). Further, he argues that qualia values must be treated as "expressions with well-defined types and relational structures" (1998:78). While it could be argued that this would have been useful for my purposes as well, for the sake of simplicity I chose to omit this information; one may assume that it is implicitly included, and to do so explicitly would not substantially affect anything proposed here or in the next chapter. Thus the qualia for the lexical item *novel* can be supplied as in (2).

(2) Qualia for *novel*[36]

    a. Constitutive: NARRATIVE

    b. Formal: BOOK, DISK

    c. Telic: READ

    d. Agentive: ARTIFACT, WRITE   Pustejovsky (1991:427)

I argued as part of the APP study that the content of qualia such as these comprises a substantial portion of the information provided by category membership. However, in order to distinguish between informativeness as I defined it and the specific portion of informativeness associated with category membership, I shall henceforth refer to the latter as "CM-informativeness."

    To summarize, for my research on APPs, I defined information as any proposition that was presented as factual and which the recipient could believe was at least probably true. However, not all information was informative, and whether a particular proposition was informative depended to some degree on the recipient. Informativeness could be achieved by any combination of a message reducing uncertainty, causing positive

---

[36] Qualia are indicated by SMALL CAPS, while linguistic items are indicated by *italics*.

cognitive effects, and increasing the manifestness of a set of assumptions. Information provided by a message was not counted towards informativeness if it was already possessed by the recipient, who could get it from a number of sources. Finally, informativeness did not seem to be reasonably quantifiable in the context of this study, so it was measured comparatively, as will be seen in the application of a further developed version of this account in the next chapter.

The connection between informativeness and relative frequency was addressed in the second paragraph of this section; it was noted that more informative phrases should occur more frequently than their less informative counterparts. This connection was used as a way of testing informativeness assessments, but in viewing the results of such tests, it is important to remember that what was being assessed when using qualia was CM-informativeness, not total informativeness. Because of this and other factors discussed shortly, the relationship to relative frequency was not as strong as one might expect. However, it could still be seen. As an example of the connection between CM-informativeness and relative frequency, consider the lexical item *novel.* Pustejovsky (1991) gives READ as its telic quale; thus it was expected that *the read novel* was not much more CM-informative than *the novel*, and certainly less CM-informative than *the unread novel*. The phrase *the read novel* was thus expected to occur rarely if at all. In fact, usage suggested that for this particular case (and generally for nouns with a telic quale of READ), this assessment was correct: a corpus query for *a/an/the read N* returned mostly phrases in which *read* was not an APP (e.g., *the read option*, a football play in which *read* is a noun), while a query for *a/an/the unread N* returned mostly phrases containing nouns with a telic quale of READ, e.g., *report*, *newspaper*, *essays*.

Because informativeness here was being assessed based on a relationship between the verb associated with an APP and the noun that occurred in the phrase with the APP, it was not possible to directly and in isolation classify APPs or their associated verbs regarding informativeness. What could be directly assessed, however, were the types of nouns that typically occurred in phrases with each APP. What proportion of them had qualia that would mean the negated form of the APP was more informative, and what proportion had qualia that would mean the non-negated form was more informative? It

was expected that APPs classified as positive typically occurred with nouns whose qualia were such that the negated form of the APP was relatively uninformative. Likewise, negative APPs would typically occur with nouns whose qualia meant that the non-negated form of the APP was relatively uninformative. In fact, this was generally the case. In 74% of the APPs examined, the top two noun collocates had qualia like those described above (COCA). In other words, positive APPs' collocates had qualia that allowed those APPs to be more informative when non-negated than when negated; that is, the non-negated APP denoted a property not supplied by category membership. Negative APPs were the opposite: the negated APP denoted a property not supplied by category membership.

Results thus suggested that informativeness assessed via noun qualia was a highly explanatory factor in the distribution of positive and negative APPs. In addition, the significance of informativeness for APPs has been recognized elsewhere. Ackerman and Goldberg make a claim very similar to the one here via their "non-redundancy constraint," which states: "If the referent of the head noun, N, implies a property P as part of its frame-semantic or encyclopedic knowledge, then an APP is not allowed to simply designate P; it must be further qualified" (1996:21). This is essentially another way of stating that APPs of the type identified by Ackerman and Goldberg's constraint are not informative, and uninformative expressions are generally less used than more informative ones.

Having defined what was meant here by informativeness and how it was assessed, I present three other factors in the APP study that I grouped with informativeness. The first of these was redundancy with resulting lexical blocking (Householder 1971; Blutner 2004). For some of the APPs tested, the negative form was essentially (at least in some senses) completely synonymous with some other unrelated form, as in the example previously given with *unbent* and *straight*.[37] In any case such as this, the synonymous forms ought to be, according to the criteria set out above, assessed as equally informative. The fact that in these cases the non-APP form was invariably more common, with the

---

[37] Note that *unbent* (in its adjectival function) and *straight* are alike in failing to distinguish between the meanings "never having been bent" and "having been bent, followed by a reversal of the bending." For *unbent*, this is because *un-* in this case can have either of the two meanings discussed earlier.

APP form often not found at all, was explained by lexical blocking. Lexical blocking, as discussed by Blutner, is a process in which "the appropriate use of a given expression formed by a relatively productive process [here, APP formation] is restricted by the existence of a more 'lexicalized' alternative to this expression" (2004:501). Aronoff (1976) and Kiparsky (1982) have shown that this process applies not only to "expressions" but also to inflectional and derivational processes such as those involved in the transformation of the simple form of a verb into an APP. As suggested by Blutner, I concluded that this blocking was the result of pragmatic factors; specifically, the use of the anomalous or less lexicalized alternative provides information by suggesting that there is a specific reason for the choice of that alternative over the commonly used item.[38] Speakers wanting to avoid this suggestion must therefore avoid the alternative; the majority of the time, this was in fact what happened.

Statistical preemption was a related factor that could either co-exist with lexical blocking or provide an alternative explanation for it. Statistical preemption, according to Boyd and Goldberg, "is an implicit inference speakers make from repeatedly hearing a formulation, B, in a context where one might have expected to hear a semantically and pragmatically related alternative formulation, A. The result is that speakers implicitly recognize that B is the appropriate formulation in such a context, and that A is not appropriate" (Boyd & Goldberg 2009:5). This could serve to further explain lexical blocking, in that speakers use the statistical rates at which they hear certain expressions to realize that one expression is, in fact, more lexicalized, and therefore use of the other would be pragmatically marked. Used as a different explanation independent of lexical blocking, statistical preemption could function as a model of how the disuse of APP forms and preference for equivalent non-APP forms is perpetuated.

Another factor related to informativeness was that of conceptual impossibility (or extreme unlikelihood): in most cases, a phrase that represents something impossible in the world, and therefore very difficult or impossible to even imagine, was not used.[39] It

---

[38] Levinson's (1987) M-principle, discussed in 2.2.3, also addresses this.
[39] One may perceive as an exception to this a phrase, such as *the unbuilt building*, that represents something that is physically impossible in the world but not conceptually impossible, and thus useful to be able to talk about.

was for this reason that, for example, that the APPs *unchewed* or *unalerted* were never found with the reversal meaning; it is impossible to unchew something or to unalert someone. Similarly, the APP *gathered* was used in the examined data only to refer to groups of things or people (e.g., *crowd*, *assembly*); a group not gathered is arguably not a group, making the idea of *an ungathered group* highly unlikely if not impossible, and in fact the APP *ungathered* was not found in my data. APPs describing conceptual impossibilities were, as a group, absent from the data, as would be expected.

3.2 Factors that affect predictions based on informativeness

From the varied aspects examined above, it was clear that there was no single factor responsible for the differences in distribution between negative and positive adjectival past participles in attributive function, but informativeness was a factor with great explanatory potential: in general, the form (negated or non-negated) of an APP that was predicted to be more informative was the form I found occurring most often in my data. This portion of my analysis was also strengthened by its general agreement with that of Ackerman and Goldberg (1996). However, as mentioned above, informativeness is highly sensitive to context, and as such there are a number of contextual and similar factors that can alter the distributions that one would otherwise expect to find. I present these here as more recent developments to the APP study.

3.2.1 Generalization of an individual-based property

Informativeness, being relative to an individual's knowledge, is necessarily relative to an individual: what is highly informative to one person might be only somewhat informative to another, and completely uninformative to yet another. This individualized nature of informativeness is further discussed in 4.1.2. Further, since an individual's knowledge is constantly changing, what is informative to someone at one time may have a different level of informativeness at another time. However, since there is not a separate corpus of each person's language use, and multiple uses of a linguistic form necessarily take place over time, looking at relative frequency unavoidably involves looking at use across many individuals over some span of time, usually at least a few years.[40] This means that, even when forms' relative frequencies are highly correlated with

---

[40] At the time of the APP study, COCA contained data from 1990 to 2010.

informativeness, what is seen in a corpus can be thought of as distribution based on an "average" of informativeness levels across language users, time, and, as discussed in the next section, situations. Thus the informativeness levels of phrases as determined for a single individual may not obviously correlate with relative frequencies of those phrases in a corpus, especially if one or more of the phrases have significantly different informativeness levels to the individual in question than to most people or an "average language user" (assuming such a thing exists).

3.2.2 Applicability of phrases

It should be obvious that one does not select a phrase to utter based on its length and informativeness alone; it must also be applicable to one's current situation and linguistic needs. Since not all situations and needs occur with equal frequency, not all phrases are applicable an equal amount of the time. Thus even if all phrases were of equal informativeness, they should not all be expected to occur with the same frequency. In fact, since CM-informativeness can be linked to unexpectedness, in the sense that something expected is more likely to be necessary for category membership or linked with a quale, it is reasonable to think that more CM-informative phrases are, by definition, applicable less often than their less CM-informative counterparts. There are, therefore, competing forces: more CM-informative phrases seem to be chosen for use in a higher proportion of the situations in which they are applicable, but less CM-informative phrases are applicable in a greater number of situations.[41] Ideally, then, one would examine relative frequency as a function of the percentage of applicable situations in which a phrase was chosen for use, but with a corpus this was simply not possible.

3.2.3 Conventionality

Another factor to be discussed is that of relative conventionality, or the degree to which particular phrases are set in ordinary usage. Idioms, perhaps the most thoroughly set phrases, are outside the scope of this discussion due to their noncompositional meaning. However, many non-idiomatic phrases display varying degrees of conventionality, and this correlates with not only the frequency with which they appear, but also the occurrence of similar, competing phrases. It is necessary, of course, to note

---

[41] This is also addressed in 4.4.2, specifically in the discussion of (6c).

that conventionality may be either the cause or the result (or perhaps even some of both) of a high frequency of occurrence – a phrase may become set in ordinary usage due to frequent use; it may come to be frequently used because, as a set phrase, it is particularly convenient or cognitively available; or both of these processes may factor in to the conventionality and frequency of a given phrase.

As an example of a conventional phrase, consider *done deal*. It is frequently non-idiomatic and typically occurs in the form *a done deal*, but is also found as *the done deal*. These two variants make up a significant percentage of phrases of the form *a/the done N*: 95% of occurrences of this form have *deal* as the N (COCA). This means that, if *a/the done deal* were not a conventional phrase, *done* would occur attributively in simple noun phrases much less frequently, changing the ratio of occurrences of *done* to *undone* in attributive position. In addition, other phrases with similar meaning, such as *a/the finished deal*, *a/the completed deal*, and *a/the finalized deal* would likely occur more often than they do.[42] Because of the conventionality of the phrase with *done*, these other phrases are rarely found: *a/the finished deal* does not occur at all in COCA, while *a/the completed deal* appears only once and *a/the finalized deal* is found twice, compared to *a/the done deal*'s 303 occurrences. Thus without the influence of the conventional phrase, words like *finished*, *completed*, and *finalized* would likely occur attributively in simple noun phrases more frequently than they now do, also changing the ratio of non-negated to negated forms for these words.

A similar situation can be seen with the phrase *a/the guided tour*, the 230 appearances of which make up 52% of occurrences of phrases of the form *a/the guided N* (COCA). Due to the conventionality of this phrase, alternatives occur much less frequently, if at all: *an/the escorted tour* is found only 5 times, while *a/the directed tour* and *a/the led tour* do not appear at all. In this case the conventional phrase seems to remove the possibility of synonymous alternatives while not making up a significant majority of the phrases using its adjective. A more dramatic case is that of a less-used but highly

---

[42] See the discussion on statistical pre-emption in 3.1.6.

conventional phrase, *a/the risen lord*[43]: its 51 occurrences make up 65% of all instances of phrases of the form *a/the risen N*, and no synonymous alternatives of the form *a/an/the Adj N* were found in COCA.

3.2.4 Use of other modifiers, such as *half-*, and incremental theme

As mentioned earlier, incremental theme was considered in the previous study, but did not seem to be highly explanatory. However, upon revisiting this notion, it becomes apparent that whether a verb takes an incremental theme or not has an effect on what types of modifiers may be used and when; this is relevant to the distribution of APPs with and without the negating prefix *un-*. Like telicity and other components of situation aspect, the notion of incremental theme applies to verbs and their arguments rather than to adjectives, and thus must be related to the verb from which an APP is derived and the themes of that verb when describing the event in question.

If an event described with a verb using an incremental theme has started, the theme in its entirety cannot reasonably be called *a/an/the unAPP [theme]*, where the APP is derived from the verb used to describe the event. For instance, once an event of mowing the lawn has commenced, *the unmowed lawn*[44] cannot be used to refer to the whole lawn, but rather only to the portion that has not yet been affected by the event. It is only when the event has not begun that the whole theme may accurately be called *a/an/the unAPP [theme]*, since once the verb's action starts taking place, at least part of the incremental theme is considered to have been "verbed." Thus it is only before an event of mowing has begun that the whole lawn can be described as *the unmowed lawn.* Similarly, most incremental themes in their entirety cannot be called *a/an/the APP [theme]* during the event. For example, during a mowing event, *the mowed lawn* cannot be used to refer to the whole lawn, but only to the portion of the lawn that has been mowed. However, while the event is in progress, one may typically use non-negating modifiers to the APP such as *half*, *partially*, and *mostly* in referring to the incremental theme, as shown in (3) below.

(3) Phrases describing incremental themes during event

       a. the half-eaten sandwich

---

[43] The word *lord* varied in capitalization in the corpus, but corpus searches were not sensitive to this; this phrase occurs in contexts such as *faith in the risen lord* (COCA).
[44] This and some other examples in this section are adapted from themes mentioned in Dowty (1991).

69

b. a partially mowed lawn

c. the mostly swept floor

A phrase such as that in (3a), with *half-*, is likely to be used when the corresponding event is roughly halfway completed; i.e., this phrase would not likely be used of a sandwich with just one bite out of it, or just one bite left of it. The modifier *partially*, as in (3b), is a bit more flexible, being applicable for a larger portion of the event than *half-*. The phrase in (3c), with *mostly*, is similar to that in (3a) in being more restricted; in this case, the modifier is used only when the even is relatively near completion. Of course, once the event is complete, the theme is described mainly as *a/an/the APP [theme]*, e.g., *the mowed lawn*, with the possibility of the occasional modifier such as *fully* or *completely*.

When an event involves a non-incremental theme, descriptions before and after the event are the same as those outlined above. However, during an event described by the verb from which an APP is derived, some non-incremental themes can be described with the APP under consideration as *a/an/the unAPP [theme]* or *a/an/the APP [theme]*, as shown below.

(4) Phrases describing non-incremental themes during event

    a. an unlicensed driver

    b. a driven car

Note that both types of description (negated and non-negated) are generally not available for the same theme: *a licensed driver* cannot be used to describe the driver during an event of licensing, and *an undriven car* cannot be used to describe the car during an event of driving. In addition, there are some non-incremental themes that cannot be described by either phrase type during the event, such as (arguably) *the cart* during an event of pushing it. Finally, modifiers such as those shown in (3) cannot be used to describe a non-incremental theme mid-event, as shown in (5).

(5)   a. *a half-licensed driver

    b. *a partially driven car

    c. *the mostly pushed cart

The unacceptability of these phrases should come as no surprise, since they describe

themes as being incrementally affected by events that are not conceived as incrementally affecting their themes.

It is clear from these patterns that the availability of different forms of APP phrases (i.e., unmodified, modified but non-negated, and negated) across time varies between events associated with incremental and non-incremental themes. Table (2) below illustrates these differences.

Type of theme associated with event ↓	Before event	During event	After event
Incremental	negated	modified[45]	unmodified; modified
Non-incremental	negated	(negated; unmodified)	unmodified; modified

Table 2: Availability of forms of APP phrases over time

As can be seen above, phrases of the type *a/an/the unAPP [theme]* are available for a larger proportion of time when describing non-incremental themes – both before and sometimes during the event, as opposed to only before the event for incremental themes. In contrast, phrases of the type *a/an/the [modifier] APP [theme]* are more available for incremental themes. Thus the type of theme typically associated with the verb from which an APP is derived is clearly a factor that has at least the potential to affect the relative frequency of negated and non-negated forms of an APP.

3.2.5 Clashes among information provided by category membership

Another contextual factor that may affect distribution of negated and non-negated APPs is the presence of clashing expectations, both established by information associated with category membership, about whether a mentioned item has or has not been, or will or will not be, the object of an action denoted by a verb from which an APP is derived. One such type of expectation arises when the item is not considered to be a member of its category once it has undergone some action that it is not particularly atypical for it to undergo. Such a situation arises with some verbs of consumption or destruction, and can be described by the statement *an APPed x is no longer an x*, as in (6).

(6) Clash with verbs of consumption or destruction

      a. an eaten sandwich is no longer a sandwich

---

[45] 'Modified' is used here to refer to modification as shown in (3), and does not include negation.

b. a shredded page is no longer a page

The statements shown above indicate that the item mentioned was once a member of the category in question but has been changed in such a way as to remove it from the category. The statement in (6a) reflects that something that qualifies as a member of the sandwich category must be uneaten, or only partially eaten, even though eating is something commonly done to sandwiches; (6b) shows that a member of the page category must be unshredded, or perhaps partially but not fully shredded, even though fully shredding pages is certainly not uncommonly done. The effect of these requirements is that phrases such as *an/the uneaten sandwich* and *an/the unshredded page* are relatively uninformative under neutral circumstances, despite the fact that they indicate the absence of an action frequently done to the noun in question, a function that would normally cause them to be relatively informative. Their low informativeness is caused by the fact that, in each case, the extra information provided by the APP can be thought of as more or less necessary for membership in the category named by the noun, and thus qualifies as information provided by category membership. Such phrases are thus expected to occur relatively rarely. In addition, phrases such as *an/the eaten sandwich* and *an/the shredded page* may, more often than would otherwise be expected, be used to refer to something that has only partially been affected by the relevant process, e.g., that which could also be called *a/the partially eaten sandwich*. This is because the alternative – that they would refer to the fully affected item – would require them to refer to a former category member as though it were a current member, e.g., calling something *a sandwich* that is no longer a member of the sandwich category. It should be noted, however, that referring to a former category member as though it were a current member is sometimes done in cases where former category membership is particularly is particularly salient and/or adds relevant information, as with *melted ice*, *paid debt*, and *corrected error*.

With some verbs of creation,[46] a situation that contrasts with the above can hold, resulting in another type of clash. This type of situation and the previously discussed one both relate to the issue of conceptual impossibility raised at the end of 3.1.6. The

---

[46] I use the term 'verbs of creation' here perhaps somewhat loosely to include any verb denoting a process that creates in an item what could be considered a defining characteristic of that item, e.g., the coiling of a spring.

particular scenario can be described by the statement *an unAPPed x is not yet an x*, as in (7).

(7) Phrases with verbs of creation

       a. an unbuilt building is not yet a building

       b. an uncoiled spring is not yet a spring

       c. an ungathered crowd is not yet a crowd

These phrases reflect expectations that any member of the building category has been (at least partially) built, any member of the spring category is coiled, and any member of the crowd category is gathered. That is, the actions referenced are required to have been done to an item in order for that item to be a member of the relevant category. The effect of these expectations is twofold: first, phrases such as *a/the built building*, *a/the coiled spring*, and *a/the gathered crowd* are relatively uninformative under neutral circumstances, since they provide already assumed information, and are thus rare. In the second effect is where one finds a clash of expectations: phrases such as *a/the unbuilt building*, *a/the uncoiled spring*, and *a/the ungathered crowd* are expected to be quite informative, since they indicate the absence of actions that are generally assumed to have occurred, and should thus occur relatively frequently. However, if one accepts that those actions, as noted above, are required for category membership (i.e., if we accept the truth of the statements in 6), these phrases refer to things that do not exist and should thus have quite limited use. This issue, though, has less to do with whether the item in question actually exists prior to the relevant action and more to do with whether it can be *conceived* as existing. For example, judging by frequency on Google, a building not yet built seems to be somewhat frequently and easily conceived as a building, perhaps because planning before the building of a structure is necessary and time-consuming, and involves the creation or procurement of related products (architectural plans, models, permits, etc.); in contrast, a crowd not yet gathered is very rarely mentioned. In sum, it can be said that phrases with negated APPs derived from verbs of creation may, in some cases, be expected to occur less frequently than informativeness assessment would indicate; the cases most affected would be those in which it is the most difficult to conceive of the item in question as existing as a category member prior to the action of

creation.

Another type of clash, quite similar to the first type discussed, may arise when an APP describes an item that is typically disposed of upon becoming describable by that APP, as with *the used tissue / paper towel / toilet paper*. These items are generally used, which would lead to the expectation that referring to them as *used* is not particularly informative under usual circumstances. However, they are also typically disposed of immediately after use, so that used versions tend not to exist in a noticeable way. Therefore, while use is expected, it is also expected that actual encountered instances of these items will most often be unused, and so identifying them as used when that is the case is, in fact, quite informative. Thus the predicted occurrence of such phrases as those above is higher than it otherwise would be. Other items with a similar effect are those that, when they can be described by a certain APP, are much less useful in life and/or in conversation, as with *the unreturned questionnaire/survey* and a number of nouns described by APPs such as *unneeded*, *unimagined*, and *unobserved*. While phrases with these negated APPs would often be otherwise expected to be relatively informative, they occur less than might be predicted due to the fact that they describe things that, for most people, there is simply not much need to describe.[47]

3.2.6 Distinction within a group

Ackerman and Goldberg note something else that I found to be true of my data in the original APP study, namely that "contrastive contexts are able to rescue APPs from unacceptability" (1996:23). In other words, a phrase that would be relatively CM-uninformative under most circumstances, and thus not likely to be used, may be much more informative, and thus useful, in a situation where it helps to distinguish some subset of a group. For example, since most grass is green, *the green grass* is typically not appreciably more informative than *the grass*. Even if most grass were not green, the former phrase would be relatively uninformative in a situation in which interlocutors can see that the grass in question is green. However, *the green grass* phrase is much more informative in a situation in which interlocutors can see that some grass is brown and

---

[47] Note, however, that some of this effect may be a result of the type of usage and users that get catalogued in a corpus; researchers may be more likely than research subjects to encounter or talk about unreturned surveys/questionnaires, but they may also be more likely to find their usage in a corpus or archive.

some is green, because it now serves to eliminate the brown grass as the object of mention. Similarly, most novels are read, so *the read novel* is relatively CM-uninformative under usual conditions, but may be highly informative if there is a restricted group of novels to which it may refer, and only one of them has been read (perhaps by some particular person). The existence of this distinctive function means that a phrase may occur more frequently than would be expected, and if the need to distinguish within a group arises often, a phrase that is relatively CM-uninformative under neutral circumstances may actually occur quite frequently.

In the same manner, a phrase that is only somewhat CM-informative under neutral circumstances will have increased informativeness when it distinguishes from a group a smaller percentage of the group than it normally would, or in other words, when it eliminates a larger subset of possibilities. When picking from all apples in the world, *a green apple* eliminates all but some subset that is probably between 25% and 75% of the group. However, if the user of this phrase has only to pick among eight red apples and two green ones, *a green apple* is more informative than in the previous situation: it eliminates all but 20% of the relevant group. If one of those green apples is subsequently removed from consideration, the phrase becomes even more informative, selecting only one apple from the group. As with the similar situation type discussed above, this may cause a phrase to occur more than would otherwise be expected, leading to an APP with higher relative frequency than predicted.

3.2.7 Creating a quantity implicature

As discussed in chapter 2, the use of an utterance with a seemingly inappropriate level of informativeness typically gives rise to some sort of implicature, or conveyance of information beyond what is strictly said. When more information than was expected is given, a frequent implicature is that all the information given is in fact necessary. When less information than was expected is given, frequent implicatures are that the speaker either does not have any more information or has the information but does not want to share it with the hearer.

Since these implicatures, as well as others, are triggered by the use of utterances that are more or less informative than what seems to be called for, they may account for some

variation between actual distributions and distributions that are expected based on informativeness. It is worth noting, however, that the adjustments in informativeness that are made for the purpose of creating implicatures do not typically involve APPs, though of course they may. For this reason, the use of quantity implicatures is best thought of as a factor that may somewhat influence distribution of linguistic items in general, but is not expected to have much of an effect on APPs in particular. For example, if A asks B how to repot his houseplant, and B knows that A is notorious for killing plants, and does not want him dealing with the houseplant, B may reply with far more information than A wants or needs, including using phrases that considerations of expected informativeness alone would suggest to be inappropriate. By providing this extra information, B creates the implicature that repotting the houseplant is a very precise, painstaking, and perhaps difficult process, thereby overwhelming and/or scaring off A so that the plant stays safe.

As mentioned above and discussed in chapter 2, implicatures can also be triggered by providing what seems like too little information; this may cause relatively uninformative utterances to be used more frequently than would otherwise be expected. For example, if A and B are in school and just had a test returned to them, A might ask B what he got on the test. If B replies with *a grade*, this is clearly a less informative answer than is appropriate in the situation and is likely to create the implicature that B does not want A to know what grade he received. B obviously has the information but does not want to share it for some reason; perhaps he is ashamed of his grade, or knows it is better than A's and does not want to make A feel unintelligent. Regardless of the exact reason, however, a phrase is used that seems to be too uninformative for the situation.

3.2.8 Attributive position

Up to this point, my discussion of factors affecting the accuracy of informativeness-based predictions has focused mainly on phrases of the type originally studied, that is, nominals with APPs in attributive position. The factors identified are, for the most part, generalizable to a variety of constructions. One factor that is form-specific, however, is the focus of this section: the relative frequencies of phrases involving APPs in attributive position seem to be more sensitive to CM-informativeness than those of phrases involving the same forms in predicative position. For example, many non-negated

adjectival participles that are rarely or never found in attributive position occur quite freely in predicative position: *an/the eaten N* and *an/the attended N* are rare, with 4 and 1 occurrences in COCA respectively, while *a/an/the N is/was eaten* and *a/an/the N is/was attended* are relatively less so, with 35 and 37 occurrences in COCA respectively. It should be noted that in predicative position these forms are typically verbal rather than adjectival, and this likely affects the degree to which they are governed by informativeness or CM-informativeness.

The exact reasons for the difference in CM-informativeness sensitivity between adjectival/attributive and verbal/predicative forms remain an area for further research. However, preliminary examination has suggested that the distinction between assertion and presupposition, and something akin to the attributive/referential distinction proposed by Donnellan (1966), may be relevant, though both can most easily be seen to apply to definite phrases, while CM-informativeness sensitivity differences are present in indefinite phrases as well. First, use of a definite nominal with an adjectival participle in prenominal position can be said to presuppose the existence of a referent of the noun with the attribute described by the adjective. In other words, using the phrase *the thrown apple* presupposes that there is an apple that has been thrown. In contrast, a predicative, verbal use of a past participle with a definite noun again presupposes the existence of a referent of the noun but now asserts, rather than presupposes, that the noun has undergone the action denoted by the participle. For example, *the apple was thrown* still presupposes the existence of an apple, but now asserts that it was thrown.

One might expect that presupposed information would face a less stringent informativeness requirement than asserted information, but this does not appear to be the case. A possible contributing factor to this has to do with a distinction similar to that which Donnellan (1966) makes with respect to attributive and referential uses of definite descriptions. Donnellan describes the attributive use as "stat[ing] something about whoever or whatever" is described and the referential use as "enabl[ing] [one's] audience to pick out whom or what he is talking about and stat[ing] something about that person or thing" (1966:285). Similarly, one can distinguish between the fact that the prenominal use of a participle often allows an addressee to "pick out" the appropriate referent from

77

among the relevant possibilities, a sort of referential use, while the predicative, verbal use of a participle more often allows one to "state something" about the referent of the relevant noun, a more attributive[48] type of use. Since information that allows one to pick out a particular member of a category over other members is closely related to what is measured by CM-informativeness, it follows that participles that are more often used attributively (by Donnellan's classification) would face higher CM-informativeness requirements. However, the analysis provided here is in its infancy, and as noted above, this area remains one with great potential for future research.

3.3 Informativeness revisited

Having presented some more recently identified factors in the APP study, and in the connection between CM-informativeness and relative frequency specifically, I now return to my original characterization of informativeness to offer an updated discussion of that as well. In this section I address several issues, some of which represent changes to the original conception, and some of which simply provide necessary elucidation.

3.3.1 Propositionality

First, I return to the topic of propositionality. At the time of the original study, the proposed characterization of informativeness was intended mainly to work with only a small subset of linguistic forms – namely APP phrases of the type studied – which could simply be converted into propositions. However, in working with informativeness in an expanded context, it has become apparent that an account that neither requires nor assumes propositional form as a requirement for informativeness is useful. In order to accomplish this with the current characterization of informativeness, I rely on the idea that a phrase, whether or not it has propositional form, can be said to 'associate' one item, action, attribute, etc. with another. I use this as a general cover term for both relationships indicated by both propositional and non-propositional forms; for example, both the constructions in (8) associate the property of having been cooked with the item denoted by *the apple*.

(8) Association

---

[48] By 'attributive' here I mean to relate this use to Donnellan's (1966) classification of attributive definite descriptions; it is an unfortunate coincidence that participles in attributive position correspond to his referential use, while participles in predicative position correspond to his attributive use.

a. the cooked apple

b. We cooked the apple

It is clear that these two constructions are not identical in either form or meaning, but in both cases there is a specific apple that can be understood to have been cooked, so both items can, in some sense, be said to have some propositional *content*, if not (in the case of 8a) propositional form. Indeed, it is the case that any association can be viewed as propositional in content, if not in form. The move from phrase conversion to the use of the concept of association thus represents a minor evolution that allows for a simpler theory by eliminating the conversion step and thereby avoiding any potential pragmatic issues resulting from the change in form. Items with non-propositional form can, therefore, be informative, and thus arises the second point on which my clarification of the propositionality requirement rests: a distinction between information and items that can be informative. Information itself can still be required to be propositional, as both constructions supply information that can be formulated as "the apple was cooked," but the items that can be said to have the property of informativeness need not themselves be propositional.

I also argue that both constructions above, to the extent that one can tell given the lack of context, are presented as factual, which I required in my original characterization of informativeness in lieu of truth. I classify as a factual presentation any presentation that does not give the addressee evidence that he should question the truth or accuracy of the associations presented. Such evidence might range from the very explicit, e.g. *It is not true that we cooked the apple*, to the subtle, such as a questioning tone of voice. Clearly, no such evidence can be detected in (8). Thus the two constructions above can be treated identically in terms of assessing the CM-informativeness of the word *cooked* in connection with the apple category; such assessment is discussed in detail in the next chapter.

3.3.2 Cognitive availability

The next task is to clarify what I mean by 'cognitive availability,' a phrase used in my characterization of informativeness but as yet undefined. To recap, my account of informativeness reads, in part, "a phrase or utterance is informative to the extent that it

increases the cognitive availability of information not already fully available." By 'cognitive availability of information,' I simply mean the ease with which that information is called to mind, or brought into conscious thought, at a particular time. This is not dissimilar to others' use of 'cognitive availability' or simply 'availability' as related to information's "activation in [one's] mind" (Linderholm et al. 2004:166), the amount of time it takes to produce a response involving a particular piece of information (Henley & Abueg 2003), or the speed with which a particular word can be retrieved (Mercer 1976). Similarly, 'lexical availability' is used, e.g. by Hernández-Muñoz et al., to describe "the ease with which a word can be generated as a member of a given category" as measured by the percentage of subjects who produce the word as a category member and the ordinal position in which it is produced (2006:730).

As should perhaps be expected, cognitive availability is not a binary, all-or-nothing property, but rather a graded one: it is not the case that one simply can or cannot call to mind a piece of information, but rather that, when one can call it to mind, the ease with which this may be done can vary. This variance depends on a number of factors, including the extent to which one has been previously exposed to the information and how relevant the information is to one's current cognitive, physical, and discourse environment. Thus information that is not already fully cognitively available can be made more available, or brought more to the forefront of one's mind, by stimulus to which it is relevant, and that stimulus can then be said to be informative according to my account. Having clarified this point, it is now possible to compare my account of informativeness to entailment-based definitions such as that used by Horn (1972). It should be clear that my characterization is much broader, in the sense that it allows a wider variety of things to be informative: one item need not entail more than another in order to be more informative, and indeed need not entail anything at all.

3.3.3 'Already available': the timing of information from category membership

The use of the phrase 'already fully available via category membership or context' in my characterization of informativeness suggests a temporal ordering of information sources; namely, that information from the category membership of entities in an utterance is received and/or processed before information provided by the associations

80

that an utterance makes. While I have no hard evidence for or against such a temporal ordering of processing, the wording I use in my characterization of informativeness is intended to capture the fact that information provided by category membership is temporally distinct from new information provided by an utterance in terms of acquisition, if not activation. The former can be considered 'prior' information in the sense that the information itself is in fact already possessed by the recipient of an utterance; the utterance simply serves to activate the information by calling to mind a particular category. In other words, the recipient already has information associated with a myriad of categories, and can create new categories with associated information as needed (see 4.1.2 for more on this), so all an utterance has to do is convey to the recipient which category (or categories) he or she should call to mind; the already-present information associated with this category then achieves varying degrees of cognitive availability, as discussed in detail in chapter 4.

3.3.4 Qualia and CM-informativeness

As detailed above in 3.1.6, I use Pustejovsky's notion of qualia to define the sorts of information provided by category membership. I showed in (2) that Pustejovsky allows for more than one quale of each type to be assigned to a lexical item; I agree with him that restricting qualia to one per type is unwarranted, but beyond that, I depart from his theories on the issue of how qualia are assigned.

In Pustejovsky's (1991) view, qualia that fully apply to the lexical item in question should be included in said item's semantics, and other qualia are not considered in relation to the item. In my view, every quale can be considered in relation to every item, though I do not propose that language users actively engage in such exhaustive consideration. Of course, given any single lexical item, most qualia will not apply at all, but there will also be a class of qualia that apply more or less well. I consider all of these qualia to be important in relation to the lexical item under consideration, and discuss how they can be managed and structured with respect to their applicability to an item in the next chapter. Thus, in my view, a lexical item can have far more qualia than Pustejovsky would likely ever propose, because I allow for graded assignment in which some qualia apply better to an item than others.

In addition, I claimed in 3.1.6 that "the content of qualia... comprises a substantial portion of the information provided by category membership," stopping short of an assertion that qualia make up *all* of the information provided by category membership. I wish to now somewhat strengthen that claim, to say that the information provided by category membership is entirely of the sort given by qualia, though not necessarily limited to exclusively the specific types of information listed in (1). That is, all information provided by category membership can be classified as constitutive, formal, telic, or agentive, but each of those classifications may contain information types beyond those listed in (1). I make this point not so much because it is critical to my notion of information that is provided by category membership, but because it allows me to conceive of the model discussed in chapter 4 as being comprehensive while having categories based on qualia, and categories whose members are qualia.

3.4 Conclusions

In this chapter I have recounted the study that motivated my characterization of informativeness and defined the concept. I also proposed a framework within which to view the portion of informativeness associated with knowledge about category membership, dubbed 'CM-informativeness.' I then discussed factors that mitigate the correspondence between relative frequency and CM-informativeness. First, CM-informativeness is related to an individual's knowledge, and as such varies somewhat from person to person, while relative frequency is a measure of use by a (usually large) population, and thus cannot be expected to reflect an individual's perceived level of CM-informativeness of an item. Second, some linguistic items are simply applicable in a greater number of situations than others, and as such occur more frequently; these tend to be those that have lower levels of CM-informativeness. In addition, conventionality also affects relative frequency. Whether a verb takes an incremental or non-incremental theme affects which modifiers can be used on its associated APP at which times, so this too can influence the proportions in which negation and other modifications are found. In some cases there are clashes between considerations of CM-informativeness, making the relationship between it and relative frequency more complex. In addition, a phrase with low CM-informativeness can have higher informativeness when used to distinguish

something within a known group. Finally, language users may choose to employ language of unexpected informativeness levels in order to create quantity implicatures.

After discussing the factors above, I then expanded and clarified some points related to my account of informativeness. First, I revised my notion of informativeness so that non-propositional items may be considered informative. Next I clarified that 'cognitive availability' refers to the ease with which one can access information at a particular point in time. I then discussed how information provided by category membership can be thought of as information that the recipient of an utterance possessed prior to the utterance. Finally, I introduced a view of qualia in which a large number of qualia may be seen as being applicable, to varying degrees, to a single lexical item.

Chapter 4. A model for assessing CM-informativeness

4.0. Introduction

Having defined informativeness, and showed how qualia may be used to predict what information can be said to be provided by category membership, I now introduce a model that uses qualia to allow for predictions to be made regarding information provided by category membership. The model I have developed utilizes two main types of categories: nominal-based categories that contain qualia, and quale-based categories that contain nominals. Though Pustejovsky (1995, 1998) provides for qualia structure of other linguistic categories as well, I concentrate specifically on nominals for the sake of simplicity and because what is shown with nominal-based categories and with nominals as category members can relatively easily be generalized to other linguistic forms. For similar reasons, the only qualia I discuss here are telic, but again, a generalization to include other types of qualia (constitutive, formal, and agentive) is quite straightforward.

The motivation for a category-based model comes from arguments for the importance of categories in cognition, particularly those made by Smith and Medin (1981), Mervis and Rosch (1981), and Lakoff (1987). Smith and Medin use the word 'concepts' to refer to the mental constructs behind the ability to categorize, or "determine that a specific instance is a member of a concept" (1981:7). They note that "without concepts, we would be overwhelmed by the sheer diversity of what we experience and unable to remember more than a minute fraction of what we encounter" (1981:1). Concepts allow us to "draw inferences about…entities," and such inferences are "activated by gaining access to [a] concept via the word that denotes it" (Smith & Medin 1981:8-9). Mervis and Rosch classify a category as something that "exists whenever two or more distinguishable objects or events are treated equivalently" (1981:89). They note that "categorization may be considered one of the most basic functions of living creatures," and without it, "an organism could not interact profitably with the infinitely distinguishable objects and events it experiences" (1981:89,94). Lakoff, following the work of Rosch (1978; 1981), supports a view of categories as prototype-based and "defined by cognitive models," which "structure thought and are used… in reasoning" (1987:9,13). He claims that "there is nothing more basic than categorization to our thought, perception, action, and speech,"

and that without this ability, "we could not function at all, either in the physical world or in our social and intellectual lives" (Lakoff 1987:5,6).

The researchers discussed above are also among those who have provided motivation for the particular nature of the categories I propose. In the next section I establish that categories in the proposed model have graded structure, vary among individuals, and can be common or ad hoc. In the following section I discuss categories based on nominals and what these can tell us about information provided by category membership. After that, in section 3, I do the same for categories based on qualia. In section 4, I discuss the relationships and interactions between categories, and synthesize the information in sections 2 and 3 to discuss the model as a whole and further outline what it can and cannot do. Section 5 delves into the differences between how the model treats informativeness and how others (particularly those subscribing to an entailment-based view) have treated it, and I conclude in section 6.

4.1 Category attributes

4.1.1 Graded structure

Both the nominal-based and quale-based categories I propose here have graded structure, which comprises three attributes: central and non-central members, fuzzy boundaries, and variation in similarity to the category among non-members (Barsalou 1983). The first two of these attributes are more important to the model than the third (with the first being the most critical), but all three will be assumed as characteristics of every category discussed in this chapter.

In categories in general, the classification of members as central and non-central, or as being in gradient distribution, reflects that some are better or more prototypical category members than others. For example, a robin is a more prototypical member of the category of birds than is a penguin; a car is a more prototypical vehicle than is an airboat. This aspect of graded structure is suggested by Wittgenstein (1953) and Austin (1961), and central and non-central members have been found in specific categories by Rosch (1973, 1975) and others. Mervis and Rosch point to a "now… growing amount of empirical evidence that all members are not equally representative of their category" (1981:95). In the context of the model proposed here, centrality reflects cognitive

availability: when one considers the base of a category, the most central members of that category are the ones that are most likely to come to mind in connection with the category base. Thus if one considers an item X, the most central members of X's category provide the information that is most present via category membership, and the more peripheral members of X's category provide the least information via category membership.

Fuzzy boundaries represent the fact that there are "unclear cases, items whose category membership is uncertain" (Barsalou 1983:211). An early category found to have fuzzy boundaries was that of furniture – there is uncertainty as to whether smaller items such as a radio are members (Barsalou 1983). Lakoff (1987) offers a somewhat different type of example with the category of rich people – how rich must one be to count as a member? Zadeh (1965) formalizes the idea of fuzzy boundaries in what he calls "fuzzy set theory," and McCawley (1981) discusses the application of this theory to language. McCloskey and Glucksberg (1978) were among the earliest researchers to identify particular category elements with uncertain membership. Mervis and Rosch note that experimental evidence has been produced of two aspects of poorly-defined category boundaries: "between-subject disagreements concerning which categories certain (poor) exemplars belong to" and "within-subject disagreement… across testing sessions" on the same task (1981:101). The concepts of fuzzy boundaries and centrality of members are both thoroughly developed by Lakoff (1987).

4.1.2 Individualization and common vs. ad hoc categories

Each category represents the way knowledge about a topic may be structured in an individual's mind. Because it is based on the modeled person's knowledge and posits actual cognitive structures and processes, the model is specific to each individual. Minor differences in membership and centrality are likely to occur, but a large amount of consistency is expected from person to person, especially within a given culture or community; such consistency has been found in subjects' ratings of a member's representativeness of its category (Mervis and Rosch 1981). This consistency is important because, in deciding what to say to an addressee, a speaker must guess at how informative his or her utterance will be for the addressee. This means that a speaker must

be able to predict, with reasonable accuracy, some aspects of how the addressee's relevant category (or categories) is arranged. However, an expectation of some variation is necessary for any account that seeks to incorporate the actual knowledge of language users and model what may be going on in the brain. All example categories given in this paper reflect the author's knowledge and posited structures.

Also, as mentioned above, each category is based on either a nominal or a quale. Because there is a theoretically infinite number of nominals, and a very large number of qualia, positing a pre-established and stored category based on each of these would be unreasonable. I therefore rely on Barsalou's notion of ad hoc categories, which are "highly specialized" categories that are "created spontaneously for use in specialized contexts" (1983:211). These are distinguished from common categories, or those that are "well established in memory," but both types of categories share the properties of graded structure (Barsalou 1983:224), so in the model presented here common and ad hoc categories are structured and function identically.

I propose that each individual has common categories based on frequently encountered nominals and qualia. The membership of one person's category based on a given item may differ from another's category based on the same item, in terms of both members vs. non-members and centrality of members. In addition, each individual has the ability to create ad hoc categories based on infrequently encountered nominals and qualia as the need arises (i.e., as an infrequently encountered nominal or quale is encountered). Like common categories, ad hoc categories are likely to differ to some extent between individuals. Further, what common categories one has, and therefore what ad hoc categories one might need to create, also varies somewhat from person to person, since individuals' differing experiences mean that which qualia and nominals are frequently encountered will differ to some degree. For example, all people can be expected to have a common category based on the nominal *food*, but the nominal *aquarium cleaning supplies* would likely require generation of an ad hoc category among non-aquarists. This can be viewed as analogous to the finding that which category level in a domain is basic, or the most fundamental, "can vary as a function of both the cultural significance of the domain and the level of expertise of the individual" (Mervis & Rosch

1981:93). Furthermore, definite nominals (discussed more in 2.2) almost always require ad hoc categories; even though the head noun and even a nominal of the same form may be frequently encountered, it is unlikely in most circumstances that the same form with the same referent is encountered frequently enough to form a common category based on it.

4.2 Nominal-based categories

4.2.1 The basic nominal-based category

The first type of category I will discuss is that which is based on a nominal; in other words, a specific nominal, hereafter referred to as the "base nominal," is the linguistic form around which the category is built, and members (and non-members) are determined according to their relationship to the concept denoted by that nominal. In this type of category, members are qualia distributed on a gradient according to the degree to which they apply to the nominal in question. These qualia and their placement within (or outside) the category can be thought of as representing an individual's encyclopedic knowledge about the concept denoted by the base nominal, but should not be conceived of as representing the 'meaning' of the nominal per se, though meaning is certainly an important aspect in determining the members and non-members of a category. Central items are proposed to be those that are most cognitively available when a category is brought to mind.[49] Recall that centrality equates to the degree to which a piece of information is considered "present" via category membership. It is important to note that the category members and non-members shown in this section are simply a representative sample of the full sets of category members and non-members, which in theory should comprise all qualia for which the author (or other modeled individual) has a mental representation. For the sake of simplicity, only telic qualia – those related to purpose and function – are considered and shown here, but it should be understood that the full set of category members also includes constitutive, formal, and agentive qualia. Because the union of the full sets of category members and non-members would contain all possible qualia of which the person whose categories are being modeled (in this case, me) is aware, it is infeasible to include them all in a graphical representation of this or any

---

[49] Central items may, but do not necessarily, correspond to definitional properties of the category base.

modeled category. Thus, even among the telic qualia alone, only some illustrative examples are provided.

Consider, for example, the nominal *apple* and the possible purposes and functions of the items to which this nominal can refer. A simplified version of *apple*'s category might look something like the illustration in figure (1).
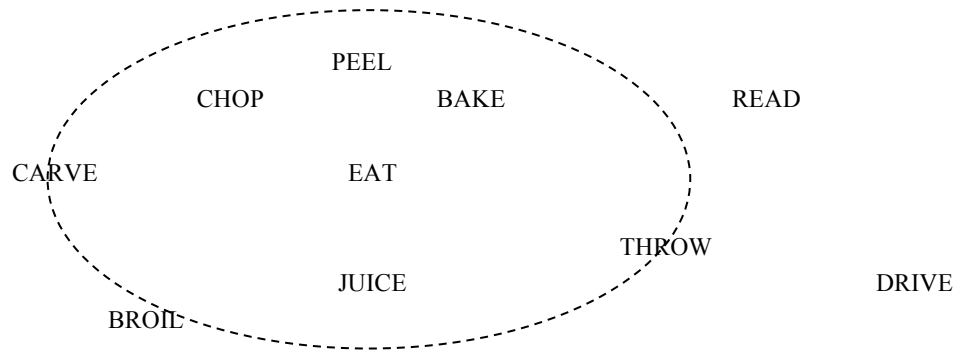

Figure 1: Category based on the nominal *apple*

Here the dashed line represents the fuzzy boundary of the category. Qualia more towards the center of the category are samples of the most central or prototypical category members, or those that apply most clearly or strongly to the nominal – what one might most immediately think of as one of the primary purposes of or things done to the nominal under usual circumstances. For example, I (as the person whose category is shown here) think that EAT is one of the primary things that gets done to apples, and it is one of the first activities called to mind when I think of them. Most people would likely agree on this, meaning that they too would have EAT as one of the central members of their category based on the nominal *apple*. Qualia nearer to but still well within the boundaries are examples of those that apply less strongly; apples may get chopped or cooked, but this is not as primary as eating them. Those on the boundaries represent the "unclear cases" for which fuzzy boundaries are stipulated. They may or may not apply at all to the nominal: apples occasionally get thrown or carved, but I may or may not think of these as actions that can be or are done to an apple. Qualia shown completely outside the boundaries represent the group (much larger than what is shown here) that do not

apply; apples are neither read nor driven.[50] Distance from the boundary represents the third category attribute noted in 1.1: non-members can vary in how far outside the category they are, based on the extent to which they are dissimilar to category members or fail to fulfill the requirements for category membership. Thus DRIVE is shown farther outside the category than READ, since I can imagine a situation in the actual world in which something could be written on an apple, and once could thus read an apple, but I cannot imagine a situation in the actual world in which one could drive an apple.[51]

Apple is given here as an example of a maximally simple nominal in order to best illustrate how a nominal-based category works, but it is worth noting that categories based on similar nominals, such as *a green apple* or *three Gala apples*, would have very similar structures. That is, with the likely exception of some formal qualia – those that distinguish an item within a larger domain – these categories would have most of the same members with the same degrees of centrality as the category based on *apple*. These similar nominals, however, are more likely to be the basis of ad hoc categories, since they are less likely to be frequently encountered. Ad hoc categories such as these would likely be formed by modifying the common category based on *apple*, a process that is further discussed in 2.2.

Food, being a hypernym of *apple,* has a similar qualia category, a simplified version of which might look something like this:
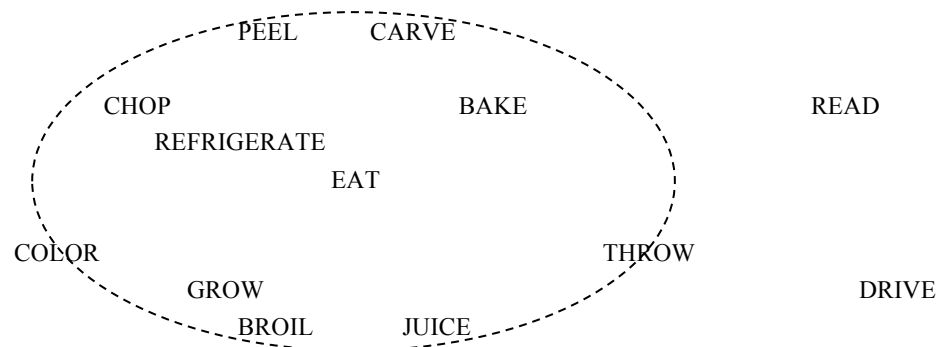
[50] 'Drive' is intended here in the sense that one drives a car (i.e., drives the instrument of driving), not in the sense that one drives a child to school (i.e., drives the contents of the instrument of driving).
[51] Note that I am confining myself to referents of *apple* that are actual apples, and excluding representations of apples. This is done for a reason similar to that which motivates confining my imagination to scenarios in the actual world. In both cases, excluding non-actual possibilities is important to the predictive power of the model, as first discussed in 2.3. In essence, I am stipulating that when one considers things that may be done to the referent(s) of a nominal, s/he considers one the most prototypical referents of that nominal and things that may be done in the actual world.

Figure 2: Category based on the nominal *food*

Again, the dashed line represents the fuzzy boundary of the category. As in figure (1), more centrally placed qualia are those that apply most clearly or strongly to the nominal; in my mind and probably every individual's, EAT is one of the primary things that gets done to food. Less central qualia apply less strongly; food may get chopped, but this is not as primary as eating it. In my perception, the majority of food is neither baked nor broiled, but baking is more commonly done than broiling, hence the former being a more central (or less peripheral) member of my *food* category. Unclear cases are shown on the boundaries; food occasionally gets thrown or colored, but I am less likely to think of these than to think of actions like eating or chopping. If this category reflected the knowledge of someone who had never heard of food coloring, COLOR would likely be farther outside the category, in the vicinity of READ and DRIVE. Those qualia well outside the boundaries represent the group that do not apply; food is neither read nor driven. As mentioned above, *food* is a hypernym of *apple*, and as a result their categories share some similarities, but, as can be seen from comparing figures (1) and (2), they are not identical.

4.2.2 Definite vs. indefinite nominals

The nominals discussed so far – *apple*, *a green apple*, *three Gala apples*, and *food* – are all indefinite. The members of their categories have been the qualia most related to the *types* of things typically denoted by the nominals, rather than to a *specific* thing denoted by the nominal. However, a category based on a definite nominal differs in a couple of ways. First, as long as the individual encountering the definite nominal knows what its referent is, the category members are the qualia most related to the *particular* thing(s) denoted by the nominal, rather than to the *types* of things typically denoted by it. Second, as mentioned in 1.3, most definite nominals will be the basis of ad hoc rather than common categories, because most definite nominals are not encountered frequently enough with the same referent for a common category to be formed.

However, as mentioned above, the ability to create a category based on a specific referent is of course contingent on knowing what that referent is, as distinct from knowing the meaning of the nominal. That is, an individual could know the meaning of

91

the apple without knowing which apple it refers to. In this case, the category based on the nominal would be ad hoc only in the sense that it was used for *the apple* rather than for *apple*; the category itself would be identical to the one based on *apple* since the individual had no further information to use to create a category. Furthermore, it is conceivable that even an individual who knows the referent of a definite nominal may still lack any further information about it, and thus be in the same position as someone who does not know the referent when it comes to building a nominal-based category. In most cases, though, knowing the referent of a nominal involves knowing more about it than what is encoded in the nominal itself. For example, if the referent is visible to an individual, he or she is likely to have information gathered from its appearance; in the case of *the apple,* this could include information about size, color, variety, freshness, location, etc. Having further information does not mean that common categories become useless; in fact, as with *apple* and *the apple* above, they remain the foundation of ad hoc categories. As discussed in relation to *apple*, *a green apple*, and *three Gala apples* in 2.1, many (if not all) ad hoc categories can be created through the modification of one or more common categories.

Consider again the definite nominal *the apple*; with knowledge of the particular apple to which it refers and further knowledge about that particular apple, an individual can form an ad hoc category by taking the common category based on *apple* and adding members, eliminating members, and changing the centrality of members as appropriate. For example, if the referent apple was known to be rather old, EAT might be made less central or even eliminated from membership and JUICE made more central, since an old apple often has an unpleasant texture but still produces palatable juice. Other adjustments might also be made, producing a category such as that in figure (3).

CHOP   PEEL

                                              READ

CARVE   JUICE   BAKE

         EAT   THROW

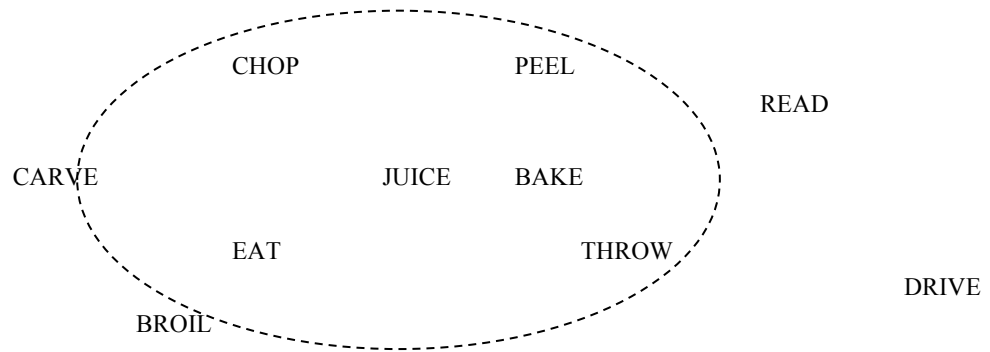   DRIVE

BROIL

Figure 3: Category based on the nominal *the apple* with known referent

Changes here from the category based on *apple* include that EAT and JUICE have essentially been switched in terms of centrality. BAKE and THROW have increased centrality, since these are more likely to be done to an apple that is unfit for eating raw. CARVE has moved slightly more towards being a non-member, since an old apple would likely be a more difficult carving medium.

Note that these adjustments of the *apple* category would basically produce the same category as the nominal *old apple*, also likely an ad hoc category produced from the *apple* category. Any ad hoc category based on a definite nominal is likely to be similar or identical to an ad hoc category based on a more descriptive but often indefinite nominal: if the referent of *the apple* was known to be large, old, and green, the ad hoc category based on *the apple* would resemble that of *a large, old, green apple*; if it was known to be old and green and to belong to one's mother, the category would resemble that of *my mother's old green apple*. Thus it is not the case that each unique nominal has a unique category (and the same can be said for qualia). Nor is it true that ad hoc categories must be created from scratch; they can often be made by modifying and/or combining existing common categories.

4.2.3 Assessment based on the model

The nominal-based type of category (as in figures (1), (2), and (3)) allows some predictions to be made. For convenience in discussing predictions, I will use the term "quale action" to refer to the action denoted by the verb of a telic quale. The first prediction that can be made is shown in (1).

93

(1)   Given a nominal-based category and a member telic quale that is relatively

        central, a phrase that associates the absence (or undoing) of the quale action with

        the base nominal is more CM-informative than a phrase that associates the

        presence (or doing) of the quale action with the base nominal.

Since a quale action (or its absence) is often associated with a base nominal through the
use of a participial adjective derived from the quale verb, in effect (1) says that, for a
central telic quale, a phrase associating the quale-derived adjective with the base nominal
is more CM-informative if the quale-derived adjective is negated than if it is not negated.

    The reasoning behind this prediction is that, if the quale is one of the primary things
that is done to the base nominal, in most circumstances it is not particularly informative
to describe the nominal as having had that done to it. Rather, this is a property that an
individual will in most cases expect of the nominal. In contrast, it is generally informative
to describe the base nominal with the negated form of the quale-derived adjective for
precisely the same reason: this attributes an unexpected property to the nominal. With
respect to figure (1), this means that a phrase that associates *food* with an adjective
derived from one of the more central qualia, such as *eaten* from EAT, is more CM-
informative when the adjective is negated than when it is non-negated, so a phrase such
as *uneaten food* is more CM-informative than *eaten food*. Recall from chapter 3 the
proposed relationship between CM-informativeness and relative frequency; more
informative items are supposed to occur more frequently than their less-informative
counterparts. In fact the former phrase does occur more often than the latter in COCA as
well as a survey of broader usage (Google).[52] As mentioned previously, all predictions
apply in unmarked circumstances, and there are, of course, exceptions, such as in
situations where expectations have been altered or there is a need to distinguish a specific
item or collection within a group.

---

[52] For each Google search, the phrase in question was searched in quotes, with multiple searches for
phrases with multiple possible determiners. I examined each result on the first 5-7 pages of results to
determine the percentage that were valid, and then calculated that percentage of the total number of results
(clicking through, when possible, to the point where the total number of results is reduced and a message is
given that "In order to show you the most relevant results, we have omitted some entries very similar to the
[#] already displayed").

The prediction in (1) can also be generalized to apply to member qualia of other (non-telic types. It will still be the case that, given a relatively central quale, a phrase that associates the nominal (or any linguistic item on which the category in question is based) with the absence of the property related to the quale is more CM-informative than one that associates the presence of the property. The main difference is that the quale is not as likely to indicate an action, or a property associated with an action having been done or not done. However, the generic term 'property' can still be used to denote whatever information the quale supplies, whether it is about an attribute, a part, the creation, or any other aspect of the entity denoted by the category base.

The second prediction that can be made with respect to nominal-based categories goes along with the previous one, and is shown in (2).

(2)   Given a nominal-based category and a member quale that is relatively peripheral, a phrase that associates the presence (or doing) of the quale action with the base nominal is more CM-informative than a phrase that associates the absence (or undoing) of the quale action with the base nominal.

Like the prediction in (1), this one can be stated in terms of quale-derived adjectives: for a peripheral telic quale, a phrase associating the quale-derived adjective with the base nominal is more CM-informative if the quale-derived adjective is not negated than if it is negated.

The reasoning for this is analogous to that of the previous prediction: one generally expects the base nominal (or indeed any nominal) not to have atypical things done to it, so describing it as non-atypical is usually far less informative than describing it as atypical on the rare occasion that it actually does have such things done to it. For example, as reflected by the fact that THROW is a peripheral quale in figure (1) above, I generally expect food not to be thrown. Therefore, describing food as *thrown* attributes to it a more unexpected property, and is thus more CM-informative, than describing it as *unthrown.* Thus *thrown food* is more CM-informative than *unthrown food*. Again considering the connection between CM-informativeness and relative frequency, we can see that this assessment is shown to be accurate by usage, as the former phrase occurs more frequently than the latter (Google).

95

As another example of the types of predictions that can be made from nominal-based categories, consider the category based on the nominal *library book,* shown in figure (4).
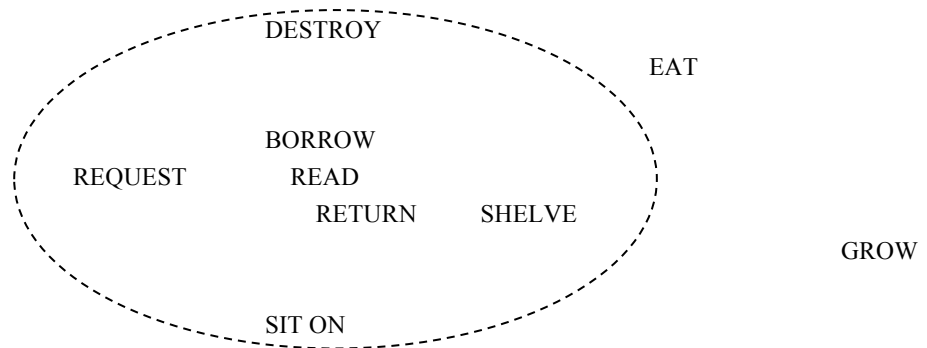

Figure 4: Category based on the nominal *library book*

For most people, RETURN is likely to be among the more central qualia in this category, since returning is one of the primary things that gets done to library books. This predicts that *a/the returned library book* should occur less often than *an/the unreturned library book*; in online usage, this is in fact the case, with the former phrase occurring 52 times and the latter 2242 (Google). Unlike returning, destroying is something that is not typically done to library books, so DESTROY would be a peripheral category member and an opposite prediction is made: *a/the destroyed library book* should occur more often than *an/the undestroyed library book*; again this is borne out, with the former phrase appearing 7 times in contrast to the latter's complete lack of occurrence (Google).

4.3 Quale-based categories

4.3.1 The basic quale-based category

The other type of category that the model uses is based on a quale (the "base quale"), with members (and non-members) determined by their relationship to that quale. For quale-based categories, members are nominals in gradient distribution according to the degree to which the base quale applies to them. Recall that a category member's centrality is representative of its cognitive availability when the category base is called to mind; thus more central members are considered to be more available via category membership. As mentioned earlier, here I will discuss only categories based on telic qualia, with the understanding that categories based on other qualia would function in the

same way. For example, consider the telic quale EAT and the items to which this action is more or less likely to be done. The category based on the telic quale EAT might resemble the following:

clay

flower

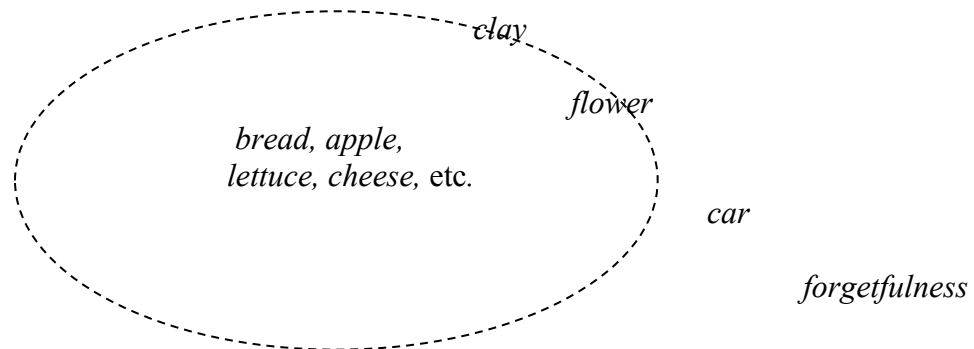*bread, apple, lettuce, cheese,* etc.

car

forgetfulness

Figure 5: Category based on the quale EAT

This illustration is set up like those in section 2; members (here nominals instead of qualia) closest to the center of the category are examples of those to which the base quale most closely applies. Peripheral members are those to which EAT applies less well, or the items which are not brought to mind as things eaten quite as easily as those in the center. On the category boundary are representatives of the items that may or may not be considered category members; they can be eaten, sometimes even without serious consequences, but are not usually eaten. Those farthest outside the category are representative of those to which it is least likely to apply (i.e., in this particular category, things that simply cannot be eaten). Note that *forgiveness* is shown farther from the category boundary than *car*; this represents the relative impossibility of eating a car – it is perhaps possible, though likely lethal, to eat at least small parts of one – versus that of eating forgiveness – completely impossible to even try due to its intangibility. As in section 2 the members and non-members shown are just a sample of larger sets that would in theory comprise all nominals, or at least all nominals for which the modeled individual has a mental representation.

Here the individual nature of the categories can be seen perhaps more easily than with the examples given in section 2: as someone who has grown and eaten edible flowers, I have *flower* as a peripheral member rather than a non-member (though it is not very far

97

within the category since I know that flowers are relatively rarely eaten, and many are not eaten at all). Someone without this experience would likely be unaware that people sometimes eat flowers, and thus have *flower* firmly outside their EAT-based category.

As another example, consider the category based on the quale READ, portrayed in figure (6).
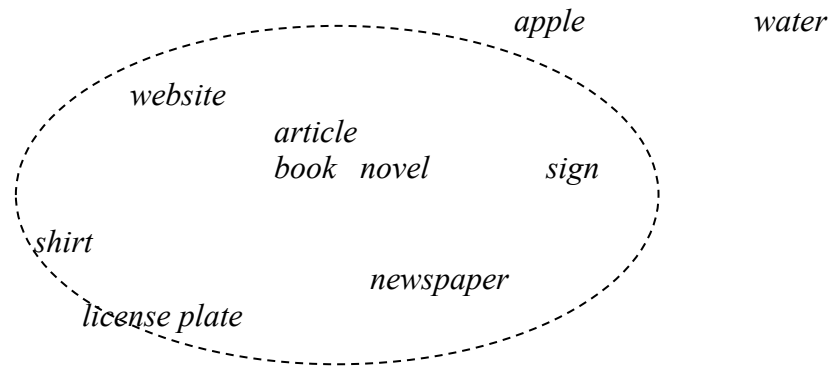

Figure 6: Category based on the quale READ

As in previous figures, central items are those to which the quale most closely applies: in my conception, books are the most likely things to be read, followed closely by articles and novels. Other items that are frequently read make up the sample less-central members shown here, such as websites, and signs. The item *apple* is shown near the border of the category as an example of something that could perhaps, in some cases, be read, but that usually is not; as mentioned in 4.2.1, one could imagine a situation in which someone writes something on an apple and it can thus be read. Finally, *water* is given far outside of the category as a sample of the large group of items that I simply cannot imagine reading under any circumstances.

4.3.2 Assessment based on the model

Like the nominal-based categories, quale-based categories also allow some predictions to be made. Analogous to the use of "base nominal" previously, I use "base quale" here to refer to the quale on which a category is based. The first prediction is given in (3).

(3)   Given a quale-based category and a member nominal that is relatively central, a phrase that associates the absence (or undoing) of the quale action with the nominal is more CM-informative than a phrase that associates the presence (or

doing) of the quale action with the nominal.

Similar to (1), this prediction essentially says that for a central nominal, a phrase associating the quale-derived adjective with the nominal is more CM-informative if the quale-derived adjective is negated than if it is not negated.

The reasoning for this can be connected to the previous category type: a nominal that is central in a quale's category is one that has that quale as central in its category. And as seen above, when a nominal has a particular quale as a central member of its category, the phrase that associates the quale-derived adjective with the base nominal is more CM-informative if that adjective is negated. Given the previously posited connection between CM-informativeness and relative frequency, with respect to figure (2) this predicts that central members such as *bread* should be described more often as *uneaten* than as *eaten,* so *the uneaten bread* should be more common than *the eaten bread*, and this is what is seen in online usage (Google). With the category in figure (2), another example similar to the one with *bread* above can be seen with *lettuce*, also a central member of the category. In both COCA and online, the negated phrase is more common than the non-negated one: *the uneaten lettuce* appears 28 times online and 26 times in COCA, while *the eaten lettuce* appears only 6 times online and 3 times in COCA.

By reasoning analogous to that outlined above for (3), the quale-based category type allows the prediction shown in (4).

(4)   Given a quale-based category and a member nominal that is relatively peripheral, a phrase that associates the presence (or doing) of the quale action with the nominal is more CM-informative than a phrase that associates the absence (or undoing) of the quale action with the nominal.

Similar to (2), this prediction in effect states that for a peripheral nominal, a phrase associating the quale-derived adjective with the nominal is more CM-informative if the quale-derived adjective is not negated than if it is negated. This too is supported by the usage data: *an/the eaten flower* occurs 60 times in Google results, while *an/the uneaten flower* does not occur at all, suggesting, via the proposed relationship between CM-informativeness and frequency, that the former phrase is in fact more CM-informative than the latter.

99

What may at first appear to be a problem with the model's predictions can be seen with the EAT-based category in figure (5). As shown above, *apple* is a central member of this category, so it is expected that *an/the uneaten apple* should be more CM-informative and therefore more common than *an/the eaten apple*. However, this is not the case; *an/the eaten apple* occurs almost twice as often in Google results as *an/the uneaten apple*. This example and others like it are important as a reminder that the model proposed here is intended only to assess one component that affects informativeness, namely the amount of information about an item that is present due to its category membership (what I have been calling CM-informativeness). The model does not assess other components of informativeness, and does not take other aspects of context into account. Therefore a distribution other than what would be expected due to the model's predictions does not mean that it has not correctly assessed CM-informativeness; it does mean that there are other factors affecting the informativeness of a given item in a given context, and this is completely expected. In the case of the phrases mentioned above, one known mitigating factor is that a very common company's logo is conventionally referred to as *the eaten apple* (though interestingly, the logo portrays an apple that has only been partially, not totally, eaten); the phrase may also be used, especially in titles of works of art, to evoke the legendary fruit from the tree of knowledge in the Garden of Eden. Despite patterns of use, the model still performs its function of determining one portion of what determines informativeness as a whole.

As another example of the types of predictions that can be made from quale-based categories, consider the category based on the quale ANSWER, shown below.
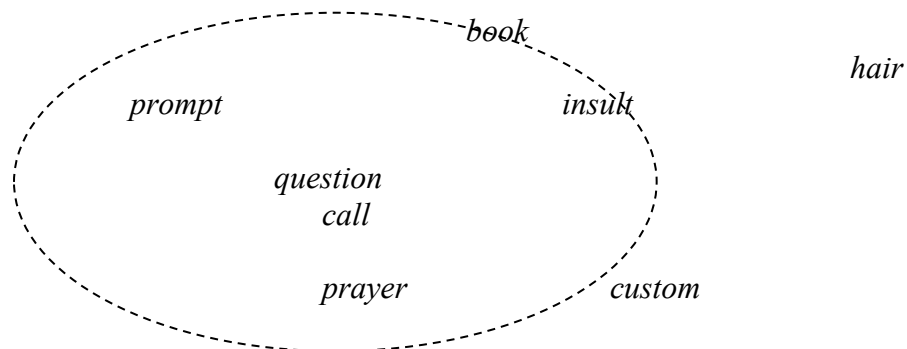

Figure 7: Category based on the quale ANSWER

A central member of this category for me, and likely for most people, is *question*. Thus a phrase that associates the adjective *unanswered* with question is more CM-informative than one that associates *answered* with *question.* As expected based on this, *an/the unanswered question* occurs more often (61 times in COCA) than *an/the answered question* (0 times). While prayers may be answered, they are perhaps not one of the most prototypical answerable things, so *prayer* may be considered a more peripheral member for many, meaning a phrase that associates *answered* with *prayer* is more CM-informative than one that associates *unanswered* with *prayer*. As predicted by CM-informativeness, the distribution of *an/the unanswered prayer* and *an/the answered prayer* is in contrast to that of the *question* phrases: the former occurs less often (4 times in COCA) than the latter (12 times).

**4.4 Category type relationships and the model as a whole**

4.4.1 Interactions between categories

Categories can be said to "interact" in two ways: they can be combined to form new categories; and one category's CM-informativeness assessments, and thus also relative frequency predictions, can be added to those of one or more other categories to yield an assessment and prediction for a more complex phrase. In determining how the CM-informativeness and relative frequency predictions gleaned from categories can interact, it is important to consider the syntactic structures involved and how they affect interpretation; this in turn shows the need for combining categories. When a noun has more than one modifying adjective in attributive position, speakers tend to process the adjective closest to the noun as being the one that became applicable to the noun first, followed by the next closest adjective, and so on, working outwards (or leftwards).[53] The basic syntactic structure with modifying attributive adjectives is as shown in (5).[54]

(5) Basic structure of a phrase with more than one adjective modifying the same noun

---

[53] The information about speakers' processing tendencies throughout this section was obtained through informal consultation with native speakers of American English.
[54] (5) shows a simple structure designed to illustrate the relationships between modifiers (specifically adjectives) and the noun they modify; I make no claims here about the presence, absence, or location of other projections.

```
          X2
         ╱╲
      Adj2   X1
            ╱╲
         Adj1    N
```
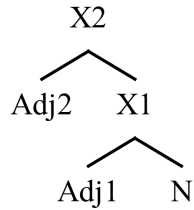
X1 and X2 are used here to avoid making any unnecessary claims about the exact nature of the constituents formed, since what is important here is only their relationships to each other. A phrase with two attributive adjectives is shown, but the same pattern of structure is thought to apply regardless of the number of modifiers in attributive position. This structure serves to explain the way such a phrase is processed, as described above. Basically, a speaker processes X1 as existing first, then Adj2 as modifying X1 to produce X2, and so on; though I make no claim about the order in which processing takes place, a clear ordering of modification is formed.

This ordering of modification becomes particularly relevant in the interaction of categories when one or more of the modifiers is a participial adjective, as are adjectives derived from quale verbs. In these cases it is assumed by hearers of this type of phrase that Adj1 came to apply to the noun first, then Adj2 came to apply to the phrase [Adj1 N], continuing similarly for any additional adjectives. With past participial adjectives in both positions, this amounts to an interpretation that the action denoted by the verb associated with Adj1 was done to the N first, followed by the action denoted by the verb associated with Adj2, and so on; this interpretation tends to occur despite the general practice of saying things in the order in which they occur, in accordance with Grice's (1975) maxim of manner. Thus the phrase [(Det) Adj2 Adj1 N] is generally not interpreted as equivalent to the phrase [(Det) Adj1 Adj2 N]. For example, when encountering the phrase *a discarded eaten apple*, hearers interpret this to mean that an apple was eaten, and then the remaining portion was discarded; *an eaten discarded apple* is, perhaps somewhat surprisingly, interpreted as an apple that was discarded and then retrieved and eaten, even though this seems much less likely to happen to an apple than it being eaten and then discarded. Likewise, *a peeled baked apple* is taken to indicate an apple that was baked then peeled, while *a baked peeled apple* indicates an apple that was

peeled before baking. A similar tendency is found when only one of the adjectives is participial: *a brown chopped apple* is more likely to be interpreted as an apple that was chopped and then turned brown, while *a chopped brown apple* is most often thought of as one that was brown before being chopped. However, interpretation according to syntactic structure is not always possible. Consider, for example, the phrase *a baked eaten apple*. Clearly an apple cannot be eaten and then baked, as a structurally-based interpretation would require. While the remains of a partially eaten apple could be baked, this interpretation also seems to be blocked, presumably by the extreme unlikelihood of one actually baking an apple core. In this type of case, hearers understand the intended meaning to be the reverse of what is structurally indicated (here, that the apple was baked and then eaten) and view the speaker as having made a mistake. Thus there is a very clear preference for interpreting the action associated with Adj1 as having happened first, followed by the action associated with Adj2, and continuing outwards through any remaining adjectives.

This type of interpretation is relevant to the interaction of categories' predictions because it affects exactly which categories are used. When Adj1 is applied to N, CM-informativeness can be assessed by simply looking at one category, the one based on N, and assessing the centrality in that category of Adj1 or a related word, such as the verb from which it is derived.[55] For example, PEEL is a somewhat peripheral member of *apple*'s category (as shown in figure (1)).

The next step, however, is not simply to look at the categories based on N and Adj2. Instead, one must consider the (almost certainly ad hoc) category based on [Adj1 N] (in our example, *peeled apple*) and assess the centrality of Adj2 or a related word. This means that the hearer of the example phrase must now create an ad hoc category based on *peeled apple* and assess the centrality of BAKE in that category, with the likely result that BAKE is a bit more central in the category based on *peeled apple* than it is in the category based on just *apple*. The ability to combine categories into a new category is not unprecedented; Smith and Medin note that "conceptual combination" allows one to

---

[55] Throughout this section I will use "related word" to mean a word with similar meaning, related via derivation (i.e., *eat* and *eaten*, but not *eat* and *uneaten*).

enlarge his or her "taxonomy of things in the world…. by combining existent [categories] into novel ones" (1981:7).

Note that, in the previous step, in theory one could consider the category based on Adj2 or a related word and introduce and assess the centrality of [Adj1 N]. I say "introduce" here because, just as it was noted in 4.1.2 to be unreasonable to expect an infinite number of nominal-based categories, it is likewise unreasonable to expect that each category with nominals as members has an infinite number of members. Thus one might allow for "ad hoc category members," or potential members that are generally not considered in conjunction with a category, but may be assessed as to membership and centrality as needed. Into this group would fall the majority of multi-word nominals. However, this does not seem warranted for two reasons. First, it is unnecessary given the established existence of ad hoc categories. Second, the centrality of a category member is intended to indicate its cognitive availability, and if an item's membership in a particular category is not considered unless necessary, this would seem to indicate very low cognitive availability of the item with respect to that category. However, as we can see with BAKE and *peeled apple* above, this would not always generate a correct assessment of CM-informativeness. Thus there seems to be no motivation for, or value to, creating ad hoc category members.

To summarize, the categories for assessing the CM-informativeness of [Adj1 N] are simple, but any subsequent assessments become more complicated as more complex ad hoc categories are introduced. CM-informativeness of a full phrase with multiple adjectives can be thought of as the sum of the CM-informativeness of each smaller phrase contained within the full phrase, and as with all CM-informativeness measures, should be thought of comparatively. Thus *a baked peeled apple* has CM-informativeness in a somewhat medium range, since the relevant centralities are somewhat peripheral (for *peeled + apple*) and somewhat central (for *baked + peeled apple*). This phrase can be thought of as more CM-informative than, for example, *a discarded eaten apple*, since EAT is central in *apple*'s category and DISCARD is likely to be central in the ad hoc *eaten apple* category, yielding low CM-informativeness for each part of the phrase and in total. Similarly, *a baked peeled apple* can be thought of as less CM-informative than, for

example, *a chopped thrown apple*, since THROW is quite peripheral in *apple*'s category, and CHOP is likely to be quite peripheral in the ad hoc category based on *thrown apple*, yielding high CM-informativeness for each phrase part and in total.

The above describes the process of assessing CM-informativeness of phrases with multiple adjectives when at least one adjective is participial. When no participial adjectives are involved, hearers may still tend to interpret phrases structurally in the absence of any cues to do otherwise, but this has little effect on the final understanding of a phrase, and is thus difficult to test. Non-participial attributive adjectives are subject to ordering restrictions that do not seem to apply to participial adjectives as much if at all; these restrictions have been explored by many[56] but are not particularly relevant to the discussion here. Assuming that structural interpretation is the norm when only non-participial adjectives are involved, the process of assessing CM-informativeness of a phrase with multiple adjectives would be essentially the same as described above. The only difference is that, since non-participial adjectives do not have associated verbs, other (non-telic) qualia would have to be used as category bases and/or members, as discussed previously.

The basic process discussed here for assessing CM-informativeness also applies to other types of phrases. As further examples I will discuss phrases that have multiple participial adjectives in predicative position, and phrases that have one attributive and one predicative participial adjective. However, this is not meant to be an exhaustive account of the types of phrases in which categories interact, but rather sufficient illustration to generalize the mechanisms of interaction to other phrase types.

With respect to this discussion, phrases with multiple predicative participial adjectives differ from their attributive counterparts primarily in their structure and resultant interpretation. Regardless of the exact structure one assumes for a phrase of the form [(Det) N copula Adj1 conjunction Adj2…], it is clear that both (or all) adjectives are relatively far from the noun, and none are given privileged position within the noun phrase. It should be unsurprising, then, that the syntactic structure here places no particular demands on the order of interpretation. Instead, hearers have a strong tendency

---

[56] See, e.g., Richards (1977), de Assis (1980), Scott (2002), and Wulff (2003).

to interpret the adjectives in the order that makes the most sense given their knowledge of the world, and a weaker tendency to interpret them in the order in which they are given. These tendencies are usually not at odds with each other, since, as was mentioned earlier, speakers generally say things in the order in which they happened. Given the phrase *the clothes were mended and ironed*, most hearers are unsure of the order in which the mending and ironing happened, though some assume the mending was first. However, *the clothes were washed and dried* clearly indicates to hearers that the washing happened before the drying, while *the clothes were dried and washed* is generally viewed as a mistake by the speaker (since clothes are generally not dried and then washed).

The result of these tendencies of interpretation is that predicative adjectives are generally assessed in the order given, with the same mechanisms discussed with respect to attributive adjectives. That is, the CM-informativeness of the pairing of Adj1 and N (for example, *washed clothes*) would be assessed first, followed by the CM-informativeness assessed through the centrality of Adj2 (or a related word) in the ad hoc category based on [Adj1 N]. For example, the ad hoc category based on *washed clothes* would be created, and the centrality of DRY in that category assessed. Then, as before, the total CM-informativeness of the phrase as a whole is the sum of the informativeness of its parts, and is a comparative measure. In the case of *the clothes were washed and dried* this would be relatively low, since WASH is likely a fairly central member of the *clothes*-based category, and DRY is likely a very central member of the ad hoc category based on *washed clothes*. An example of a more CM-informative phrase of the same form would be *the clothes were torn and sold*, since TEAR would be a relatively peripheral member of the *clothes*-based category, and SELL would be a very peripheral member of the ad hoc category based on *torn clothes*.

When a phrase has one attributive and one predicative participial adjective, i.e. [(Det) Adj1 N copula Adj2], the basic mechanisms of CM-informativeness assessment are again the same as discussed above. In this case, hearers tend to interpret the attributive adjective(s) as having applied first. Thus *the washed apple was eaten* is interpreted as meaning that the apple was washed and then eaten, and has its CM-informativeness assessed with the categories based on *apple* and *washed apple*. Though I have not tested

larger phrases of this type, given this information and the ordering of phrase types discussed above, one could hypothesize about the interpretation of a more complex phrase such as [(Det) Adj1 Adj2 N copula Adj3 Adj4]. Attributive adjectives would be interpreted as applying first, starting with the one closest to the noun, then predicative adjectives in the order given; i.e. Adj2, Adj1, Adj3, Adj4. In this case, several ad hoc categories would be created, based on [Adj2 N], [Adj1 Adj2 N], and [Adj3 Adj1 Adj2 N].

4.4.2 A generalized account of the model

The model, as shown above, assesses one factor of the degree to which specific phrases or co-occurrences of words are informative, and as such can to some extent predict relative distributions of phrases of some types as well. Use of the model is illustrated above mainly using one particular type of phrase, in which a determiner is followed by a participial adjective and then a nominal. As discussed in 3.2.8, adjective use in attributive position seems to be more sensitive to considerations of CM-informativeness, and possibly informativeness in general, than adjective use in predicative position. This does not mean the model will not work as well for assessing the CM-informativeness of predicative expressions, but it does mean that situations such as the one discussed with *an/the eaten/uneaten apple*, where the actual relative frequency and model-based prediction do not appear to match, are more likely to be encountered with predicative expressions than with attributive ones.

Though it is discussed primarily with respect to nominals with participial adjectives, the model presented in this chapter is useful for a wider variety of expressions than just nominals of this type, and may be generalized to a variety of linguistic forms. Essentially what it assesses is the degree to which specific co-occurrences of words are CM-informative, or the degree to which the information that a phrase provides about an item is not provided simply by the recipient knowing the category of which the item is a member. For prediction of relative frequency, the model works best on expressions of types that are highly constrained by informativeness, and by CM-informativeness in particular, such as noun phrases with attributive adjectives; the model relies on the fact that there is a general preference, ceteris paribus, for more densely informative

expressions – i.e., given two equally informative expressions, the shorter one is preferable, and given two expressions of equal length, the more informative one is generally preferable. For simply assessing CM-informativeness, the model works on a wide variety of expressions: any construction that creates an association between almost any two linguistic items (e.g., nominal and adjective; verb and subject or object) is assessable. The model works by examining the relationship between a category base and some category member; the category member's position in the base's category determines whether the pairing of the base and member is more or less CM-informative. For some expressions, CM-informativeness is predictive of relative frequency, with more CM-informative expressions occurring more often than less informative ones.

A general account of the model's assessment mechanism can be formulated as follows:

(6)   Given an association of the base of a category (or its derivative) B with a category member or non-member, X:

a. If X is a relatively central member of B's category, the association is less CM-informative than if X were more peripheral.

b. If X is a peripheral member of B's category, the association is more CM-informative than if X were more central.

c. If X is a non-member of B's category, the association has the potential to be even more CM-informative than if X were peripheral.

d. Negating either B or X reverses these relations.

Shown in (6) are the core assessments of CM-informativeness that are the main purpose of the model and apply to any expression of the type stated. Note that CM-informativeness measurements are given comparatively: a pairing of X and B is *more* CM-informative if X is peripheral than if X is central, and has the *most* potential for CM-informativeness if X is a non-member of the category based on B. However, if either B or X (but not both) is negated, the CM-informativeness measures are reversed, so that, e.g., a pairing of *un*X and B is more CM-informative if X is a central member of B's category than if it is a peripheral member.

In (7) are the predictions regarding relative frequency that result from the assessments

in (6); as noted previously, these predictions are sensitive to type of expression, and it is important to remember that not finding the expected distribution does not mean that any of the CM-informativeness assessments are inaccurate. Examples of nominals with attributive participial adjectives are shown here to illustrate the predictions, since this phrase type has been found to have relative frequency that is particularly sensitive to CM-informativeness.

(7)[57]   Given the base B of a category and some category member or non-member X:

a. If X is a relatively central member of B's category, the X B association usually occurs with either X or B negated.

an/the unread novel	12485
a/the read novel	329

b. If X is a peripheral member of B's category, the X B association usually occurs with neither X nor B negated.

a/the thrown apple	5505
an/the unthrown apple	0

c. If X is not a member of B's category, situations rarely if ever arise in which the X B association could be truthfully related to the actual world, so the association occurs rarely, if ever.

an/the employed peach	0
an/the unemployed peach	0

The prediction in (7a) is related to the CM-informativeness assessment in (6a); negating either X or B increases the otherwise relatively low CM-informativeness of an association of B with a central X. Similarly, the prediction in (7b) arises from the assessment in (6b); the association referenced is more CM-informative when not negated, so it is most likely to be non-negated. Given these patterns, one might expect that an association of the type referenced in (6c) would almost certainly be non-negated, since it is potentially highly CM-informative in that form. However, the difference between a peripheral category member and a non-member (as in 6b and 6c) is qualitatively different than that between a central and a peripheral category member (as in 6a and 6b). That X is

---

[57] Frequencies given here are from Google searches as described in footnote 49.

a non-member of the category based on B means that X and B are rarely, if ever, associated with each other in the world. As such, the need to associate them with each other in language also arises rarely, if at all. When necessary, such a linguistic association can be made in a grammatically and pragmatically acceptable way, and is indeed highly informative, as noted in (6c); however, as indicated (and shown by example) in (7c), the association does not occur frequently in either negated or non-negated form.

4.5 Informativeness in the model vs. in other theories

It should be evident from the discussion and illustrations up to this point that there are several differences between how my proposed model treats informativeness and how it is treated by most other theorists. In this section I will discuss in particular the differences between my view of informativeness, as implemented by the model, and the views discussed in chapter 2.

To begin, I repeat my characterization of informativeness: the extent to which a phrase or utterance increases the cognitive availability of information not fully present via category membership or context, and thereby reduces the number and/or likelihood of possibilities for the event, state, or entity to which it refers. In explicitly considering what information the recipient already has, my account is similar to those of relevance and information theories as well as some philosophical accounts, though Dretske (1981), for example, distinguishes between the information generated by an event and the information contained in a message about that event, rather than between the information contained in a message and the net information gain of the message's recipient. Though Grice does not explicitly note that the recipient's prior information should be taken into account, his injunction to be "as informative as is required (for the current purposes of the exchange)" does not preclude that he may have considered "the current purposes of the exchange" to include not giving the recipient information he or she already has, as I assumed in my discussion of quantity implicatures (Grice 1975:45). Horn (1972; 1984) does not directly address the issue of a recipient's prior information with reference to informativeness either, but his (1972) account of presupposition indicates at least an awareness of the importance of the recipient's cognitive state.

Though my theory is similar to others in considering a recipient's previous

knowledge, it is unique in explicitly taking into account information that can be said to come from the recipient's cognitive context, but to be prompted by an utterance – that is, structured information that comes from category membership. To recap the explanation in 3.3.3, upon hearing (for example) a nominal, the recipient immediately has cognitively available the information associated with the qualia category based on that nominal, with more central members of that category being more cognitively available. Hence information provided by category membership can be thought of as a recipient's previous knowledge, since it comprises only information that the recipient had prior to the utterance in question, but it can also be thought of as co-occurring with the utterance to some degree, since it is only "activated," or made cognitively available, upon the recipient's hearing (or reading, etc.) of the utterance.

Other theories vary in how explicit they are about where any 'prior information' that may be considered comes from. Relevance theory is most like my own in this respect, and it is this theory from which I have adapted my rather broad notion of context. Sperber and Wilson define context as "a subset of the hearer's assumptions about the world;" basically, anything the hearer believes may be used in interpreting an utterance (1995:15). This definition, while perhaps a bit vague, allows for some flexibility: it admits, I would argue, the possibilities of categories such as I have proposed being a part of context. The other theories discussed here are even vaguer, in that they do not explicitly address what sort of prior information a recipient might use, where that information might originate, or how it might be structured or accessed.

Another way to differentiate among theories is based on whether they portray information's presence as binary – a given piece of information is either present or not present, available or not available – or as gradient. In my account, information is conceived of as primarily gradient, though there is still a binary aspect to the theory in the classification of items as members or non-members, a distinction that does affect distribution as described in section 4.4.2. Of the other theories being discussed here, relevance theory is most like my own in this respect: it clearly allows for degrees of cognitive availability of information. Some philosophical views seem to sidestep this issue, as with Dretske's (1981) distinction between information and knowledge, the latter

of which is a belief based on (true, informational) evidence. It is not clear whether Dretske allows for the possibility of knowledge being associated with different degrees of belief. Information theory allows for varying degrees of uncertainty; uncertainty is purely a function of the number of possibilities that exist, but it does take into account that one may know some possibilities to be more likely than others. When it comes to varying availability or certainty of information, Grice (1975) and Horn (1972) say nothing to indicate support for or rejection of either side; it is worth noting, however, that both discuss the cancellability of implicatures, and Horn additionally discusses the fact that they may be strengthened into assertions. This suggests that both Grice and Horn are amenable to a theory that places importance on the idea of levels of certainty or availability of knowledge.

Finally, in differentiating between theories it is important to reiterate what I have strived to highlight throughout my discussion: the proposed model is intended to allow for the assessment of what information is present via category membership and how available it is. The model does not account for the other important information noted in my characterization of informativeness, namely everything that comes from any other aspect of context, hence my use of "CM-informativeness" to designate the measure provided by the model. This distinction is one major difference between my theory and the theories of others, who have largely treated informativeness as a monolithic whole rather than breaking down the assessment of already-present information by source. However, this focus has allowed me to be more explicit about the mechanisms behind and structure of this information source, while, as mentioned above, there has been a tendency towards vagueness among other theorists. In addition, I have been able to show that comparative levels of CM-informativeness play a role in the relative frequencies of some constructions; to my knowledge, no other informativeness assessment method has been used in this way. My account thus has a different focus than other linguistic or non-linguistic accounts; it can be said to be superior to other accounts in its ability to assess a particular aspect of informativeness, its provision of a nuanced measure thereof, and its applicability to an area in which informativeness was previously not well developed.

4.6 Conclusions

In order to assess the information discussed in a formal and individualized way, in this chapter I introduced a model that uses the notion of qualia as well as gradient, fuzzy-bounded categories. I discussed telic qualia-based categories with nominals as members, and nominal-based categories with qualia as members (and though I discussed only telic qualia members, it was noted that all qualia types exist as members). In addition, there are other categories based on other qualia types and linguistic items. The model is useful for assessing informativeness, and to the extent that it does that, can also be used to predict relative distribution in unmarked contexts. However, in its current form it is only intended to assess the aspect of informativeness related to information provided by category membership, or CM-informativeness, so in many cases numerous other factors relating to informativeness and distribution may be at work. This is not a failure of the model, but rather a necessary result of its somewhat narrow focus, which may be broadened in the future.

Following my explanation of the model, I discussed how my theory in general differs from others' theories; the focus on CM-informativeness is the most obvious difference, and the way I structure the information provided by category membership – using categories and qualia – is unique. Another prominent divergence from all but relevance theory is my treatment of information as something that can have degrees of cognitive availability, rather than simply being present or absent. Finally, I have strived to be as explicit as possible in my characterization of informativeness and of how CM-informativeness may be assessed; this has led me to also note that all of an utterance recipient's prior knowledge (category membership-related and otherwise) should be considered in assessing informativeness as a whole.

Chapter 5. Conclusion

5.0 Introduction

This work represents an attempt to address the particular needs of a study and to enrich the linguistic conception and study of informativeness by developing a new characterization of informativeness and a model of how cognition may function to assess one aspect of informativeness so characterized. In motivating, developing, and explaining my account and model, I have accomplished a number of things. I have shed light on a relatively unexamined phenomenon of distribution – that of nominals with APPs in prenominal position. I have also highlighted some of the similarities and differences between various conceptions of informativeness, both within and beyond linguistics. In addition, I have provided a broad but explicit characterization that allows for a more nuanced view of informativeness than was present in previous theories and that helps to explain the distribution of APPs. Lastly, I have discussed cognitive aspects of the calculation and application of informativeness and created a model for how a previously under-studied portion of informativeness may be assessed. This allows for further insight into an important aspect of language and cognition that has not previously been studied in this way.

I conclude here by first reviewing what I have proposed throughout this work. Then, in 5.2, I outline some limitations of the current research and suggest remedies for some of these along with other directions for future research. Finally, in 5.3 I provide some concluding remarks.

5.1 Review of proposals

Following an examination in chapter 2 of informativeness in information theory, philosophy, and the work of Grice, Horn, Levinson, and Sperber & Wilson, I first provided further motivation for my proposals by recounting my 2010 study of phrases with prenominal adjectival past participles in chapter 3. This study showed that, in the phrase type studied, some APPs occur more frequently with the negating prefix *un-* than without it. Examination of factors related to semantic fields, situation aspect, incremental theme, and word origins revealed many possible but weak correlations to relative frequency, but no highly explanatory element. Informativeness, however, was found to be

much more explanatory than any of the previously examined factors; APPs that were most often informative in negated form were those that occurred most often in negated form, which conformed to the expectation that more informative forms are more likely to be used.

Given the discovery of informativeness as an important explanatory factor, a satisfactory characterization of it was needed. Definitions formulated previously by others were inadequate, so I developed my own account, by which informativeness is a measure of the extent to which an utterance increases the cognitive availability of information that is not already fully available. Information was said to be already available if it was present anywhere in the recipient's cognitive context; this included information that came from knowledge about category membership, a unique attribute of my account of informativeness. I then focused on how to assess that specific aspect of informativeness, developing a structure based on qualia. Upon utilizing this more satisfactory account, and the new form of assessment, it was found that there were a number of mitigating factors that affected the relationship between CM-informativeness and relative frequency of APPs in the phrase type studied. I therefore discussed these factors and their effects before moving on to update and expand on my (2010) characterization of informativeness.

In again discussing my conception of informativeness, I revised it to apply to both propositional and non-propositional items and clarified what was meant by 'cognitive availability.' I discussed that category membership-related information could be considered to be already possessed by a recipient and simply activated by an utterance, hence it being characterized as information that was already present prior to the utterance. Finally, I noted that, as I use them, qualia can apply more or less well to a linguistic item, and an item can therefore have many qualia in gradient applicability.

In chapter 4 I proposed a model of how information may be arranged cognitively to allow for assessment of CM-informativeness. The model uses categories that have graded structure, are to some extent individualized, and can be common or ad hoc. Each category is based on a quale or a linguistic item. In a quale-based category, the members are linguistic items to which that quale applies, and the most central members are the ones to

which the quale best (or most closely) applies. In a linguistic item-based category, the members are qualia that apply to the base item, again arranged with the most central having the closest relationship to the base. The categories I illustrated in chapter 4 used telic qualia and nominals, but all types of qualia and linguistic items can be used in the same manner.

Using these categories, one can comparatively assess the CM-informativeness of any construction that associates a base (or related word) with a category member (or related word). When a base and a relatively central member are associated, the result is less CM-informative than the association of the base with a relatively peripheral member. Association between a base and a non-member is unlikely to occur, but when it does, is the most informative type of association.

Following my explanation of how the model and its assessment mechanism work, I discussed category relationships and interactions. I showed that, if a category based on a quale Q has a nominal or other linguistic item n as a central member, then the category based on n will have Q as a central member. Likewise, if Q has n as a peripheral member, n will have Q as a peripheral member, and if n is a non-member of Q, Q will be a non-member of n. I then showed how multiple categories can be combined to form new categories, and how CM-informativeness assessments work when multiple categories are involved. I discussed that the order in which categories are used is important, and this order varies depending on the syntactic structure of the expression being assessed and on whether adjectives are participial or not. Prenominal adjectives are assessed starting with the one closest to the noun and working outward (right to left), while predicative adjectives are assessed from left to right. When both prenominal and predicative adjectives are present, the prenominal ones are assessed first. There is little evidence that this ordering matters much when adjectives are not participial, but when they are participial, recipients tend to view the ones assessed first as having come to apply to the noun first, meaning that the action denoted by the related verb happened before any other denoted actions. This yields differing interpretations of, for example, the phrases *a peeled baked apple* and *a baked peeled apple.*

I conclude chapter 4 by outlining how my conception of informativeness, and the

116

model that goes along with it, are different from those of the theories discussed in chapter 2. The most obvious difference is the focus on CM-informativeness and the provision of a model for assessing that aspect of informativeness. I conform with information and relevance theories in explicitly considering a recipient's prior information. In other details, my conception of informativeness is most similar to that of relevance theory: both define a relatively broad context from which previous information may come and allow for information to have degrees of availability rather than simply being present or absent. I have thus presented a theory of informativeness that is influenced by, and to some degree in concord with, other theories while also making use of the unique conception of CM-informativeness.

5.2 Limitations and future research

The primary limitations of the research presented here involve its scope: there are a number of topics on which further inquiry is warranted. Several of these relate to the initial APP study, because it was the motivation for, rather than the focus of, my theory of informativeness. First, the study was limited in that 204 APPs were examined in one corpus (plus some Google results when the corpus was insufficient); increasing the number of APPs and/or corpora in a future study may provide more robust results. In addition, some of the factors investigated in the study may benefit from further research; in particular, situation aspect was difficult to assess, and a survey of native speakers' intuitions on issues such as telicity and durativity is planned as part of future work. Another area related to the APP study in which there is clearly more research to be done is the distinction in CM-informativeness sensitivity between adjectival/attributive and verbal/predicative forms as discussed in 3.2.8.

Concerning the proposed model, an obvious limitation is the lack of testing it has undergone; aside from the varyingly robust connection between CM-informativeness and relative frequency, it has so far been unclear exactly how to test the functionality of what I have proposed. In future research I would like to figure out how to test category contents (i.e., which items are members, how central they are, and which items are non-members), which would also allow me to test how an individual's categories might change over time or in different contexts. In addition, it may be helpful to determine

117

whether, for a particular individual, a particular category is common or ad hoc; this should be relatively straightforward given previous work that has been done on the common / ad hoc distinction. I believe there is also the potential for further explication on category interactions and on the differences between categories based on definite and indefinite expressions.

Finally, another limitation is the fact that, in its current form, the model only works with literal language. This is of particular concern to me because a great deal of language is non-literal, so the model is unequipped to deal with a significant portion of language as it is actually used. Thus I provide here a basic outline for how I propose to expand the model in future work. As in chapter 4, I confine my discussion to nominal- and telic quale-based categories, with the understanding that generalization to other category types would be relatively straightforward.

For a nominal-based category, one must consider what other nominal the base can be metaphorically associated with. Then the qualia of that metaphorically associated nominal appear with the appropriate degree of centrality (based on robustness of metaphor and prototypicality aside from metaphor) in the base nominal's category on a metaphorical plane. It may be useful to think of this in a 3-dimensional way, where a particular metaphorical extension is represented by a particular plane on which metaphorically associated qualia appear. Metaphorical associations only go one direction; for example, with the metaphor 'life is a journey,' the qualia associated with *journey* appear in *life*'s category, but the qualia associated with *life* do not appear in *journey*'s category. Predictions based on the category work basically the same as discussed with reference to non-metaphorical categories.

For a quale-based category, the system is basically the same: nominals to which the base quale is metaphorically extended appear with the appropriate degree of centrality (as described above) in the base quale's category on a metaphorical plane. As with nominal-based categories, metaphorical associations are one-way; given the metaphor 'seeing is touching,' the nominals associated with the quale SEE appear in TOUCH's category, but the nominals associated with TOUCH do not appear in SEE's category. Thus both category types can be extended to work with metaphorical language use in addition to literal use.

5.3 Concluding remarks

In this work I provided motivation for a new conception of informativeness by outlining the previous ways it has been used and showing that it is an explanatory factor in the distribution of negated and non-negated adjectival past participles in attributive position. I defined informativeness as the extent to which an utterance increases the cognitive availability of information that was not already fully available to the recipient. I focused on the assessment of one part of the information already available to the recipient: the information that a recipient has when he or she knows the category membership of an item. Informativeness related to category membership, or CM-informativeness, was shown to be correlated to APP frequency. I then proposed a model for assessing CM-informativeness. The model uses categories based on qualia and linguistic items; a phrase associating a category base and a central member is less CM-informative than one associating a category base and a peripheral member. The account and model I proposed allow for the assessment of an aspect of informativeness that was previously largely ignored, and provide a glimpse into some of what may be going on cognitively when language users assess informativeness. They are intended to provide further elaboration on the concept of informativeness so that it may be useful in a wider variety of applications.

References

Ackerman, Farrell & Adele E. Goldberg. 1996. Constraints on adjectival past participles. In Adele E. Goldberg (ed.), *Conceptual structure, discourse and language*, 17-30. Stanford, CA: CSLI Publications.

Aftab, Cheung, Kim, Thakkar & Yeddanapudi. 2001. Information theory: Information theory and the digital age. mit.edu/6.933/www/Fall2001/Shannon2.pdf (September 2011).

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Ash, Robert. 1965. *Information theory*. New York: Interscience Publishers.

Assis, Ana Antônia de. 1980. *The order of prenominal adjectivals in unambiguous unbroken strings in English*. Universidade Federal de Santa Catarina dissertation.

Atlas, Jay David & Stephen C. Levinson. 1981. It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In Peter Cole (ed.), *Radical pragmatics*, 1-62. New York: Academic Press.

Austin, John L. 1961. *Philosophical papers.* Oxford: Oxford University Press.

Bagassi, Maria, D'Addario, Marco, Macchi, Laura & Sala, Valentina. 2009. Children's acceptance of underinformative sentences: The case of *some* as a determiner. *Thinking & Reasoning* 15(2), 211-235.

Barsalou, Lawrence W. 1983. Ad hoc categories. *Memory & Cognition* 11(3). 211-227.

Blutner, Reinhard. 2004. Pragmatics and the lexicon. In Laurence R. Horn and Gregory Ward (eds.), *The handbook of pragmatics*, 488-514. Malden, MA: Blackwell.

Bruner, Jerome S. & O'Dowd, Donald. 1958. A note on the informativeness of parts of words. *Language and Speech* 1(2), 98-101.

Calhoun, Sasha. 2010. How does informativeness affect prosodic prominence? *Language and Cognitive Processes* 25(7-9), 1099-1140.

Carston, Robyn. 2002. *Thoughts and utterances: The pragmatics of explicit communication*. Malden, MA: Blackwell Publishing.

Clarke, James & Lapata, Mirella. 2010. Discourse constraints for document compression. *Computational Linguistics* 36(3), 411-441.

Davies, Catherine & Katsos, Napoleon. 2010. Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua* 120, 1956-1972.

Dennett, Daniel C. 1969. *Content and consciousness*. New York: Humanities Press.

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75(3). 281-304.

Dowty, David. 1987. Aktionsarten, NP semantics, and the structure of events. Paper presented at the joint Association for Symbolic Logic/Linguistic Society of America Conference on Logic and Natural Language, Stanford University.

Dowty, David. 1991. Thematic proto-roles and argument selection. Language 67(3): 547-619.

Dretske, Fred I. 1981. *Knowledge and the flow of information*. Cambridge, MA: The MIT Press.

Duym, Suzanne van. 2010. The distribution of English adjectival past participles. *LSO*

Working Papers in Linguistics (Proceedings of WIGL 2010) 8. 1-22.

Fogerty, Daniel & Humes, Larry E. 2010. Perceptual contributions to monosyllabic word intelligibility: Segmental, lexical, and noise replacement factors. *Journal of the Acoustical Society of America* 128(5), 3114-3125.

Geurts, Bart. 2010. *Quantity implicatures.* Cambridge, UK: Cambridge University Press.

Giora, Rachel 1988. On the informativeness requirement. *Journal of Pragmatics* 12. 547-565.

Green, Georgia M. 1996. *Pragmatics and natural language understanding.* Mahwah, NJ: Lawrence Erlbaum Associates.

Grice, H. P. 1957. Meaning. *The Philosophical Review* 66(3), 377-388.

Grice, H. P. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics, volume 3: Speech acts*, 41-58. New York: Academic Press.

Grice, H. Paul. 1978. Further notes on logic and conversation. In Peter Cole (ed.), *Syntax and semantics, volume 9: Pragmatics*, 113-128. New York: Academic Press.

Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grice, H. P. 1991. Utterer's meaning, sentence-meaning, and word-meaning. In Steven Davis (ed.), *Pragmatics: A reader*, 65-76. New York: Oxford University Press.

Grimshaw, Jane & Sten Vikner. 1993. Obligatory adjuncts and the structure of events. In Eric Reuland & Werner Abraham (eds.), *Knowledge and Language, vol. II, Lexical and conceptual structure*, 143-155. Dordrecht, Netherlands: Kluwer Academic Press.

Gundel, Jeanette Kohn & Ann E. Mulkern. 1998. Quantity implicatures in reference understanding. *Pragmatics & Cognition* 6(1/2). 21-45.

Henley, Nancy M. & Joselito Abueg. 2003. A review and synthesis of research on comprehension of the masculine as a generic form in English. *Estudios de Sociolingüística* 4(2). 427-454.

Hernández-Muñoz, Natividad, Cristina Izura & Andrew W. Ellis. 2006. Cognitive aspects of lexical availability. *European Journal of Cognitive Psychology*. 18(5). 730-755.

Hinrichs, Erhard. 1985. A compositional semantics for aktionsarten and NP reference in English. Columbus: Ohio State University dissertation.

Hirschberg, Julia Linn Bell. 1985. *A theory of scalar implicature.* Philadelphia: University of Pennsylvania dissertation.

Horn, Laurence Robert. 1972. *On the semantic properties of logical operators in English*. Los Angeles: University of California dissertation.

Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (ed.), *Meaning, form, and use in context: Linguistic applications (GURT '84)*, 11-42. Washington, DC: Georgetown University Press.

Horn, Laurence R. 2004. Implicature. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 3-28. Malden, MA: Blackwell Publishing Ltd.

Householder, Fred W. 1971. *Linguistic speculations*. London: Cambridge University Press.

Kiparsky, Paul. 1982. Word-formation and the lexicon. In F. Ingeman (ed.), *Proceedings of the 1982 Mid-America Linguistic Conference*. Lawrence, KS: University of

Kansas.

Krifka, Manfred. 1987. Nominal reference and temporal constitution: Towards a semantics of quantity. (FNS-Bericht 17). Tubingen: Forschungsstelle fur naturliche Systeme, Universitat Tubingen.

Krifka, Manfred. 1989. Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aktionsarten. Munich: Wilhelm Fink.

Lakoff, George. 1970. *Irregularity in syntax*. New York: Holt, Rinehart and Winston, Inc.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, Robin Tolmach. 2009. Conversational logic. In Jef Verschueren & Jan-Ola Östman (eds.), *Key notions for pragmatics*, 102-113. Amsterdam: John Benjamins Publishing Company.

Leech, Geoffrey N. and Jennifer Thomas. 1988. Pragmatics: The state of the art. *Lancaster Papers in Linguistics* 48.

Leeuwen, van Marieke, Berg, van den Stéphanie, Hoekstra, Rosa A. & Boomsma, Dorret I. 2009. The genetic and environmental structure of verbal and visuospatial memory in young adults and children. *Neuropsychology* 23(6), 792– 802.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge, UK: Cambridge University Press.

Levinson, Stephen C. 1987. Minimization and conversational inference. In Asa Kasher (ed.), *Pragmatics: Critical concepts, volume IV: Presupposition, implicature, and indirect speech acts*, 61-129. London: Routledge.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature.* Cambridge, MA: MIT Press.

Linderholm, Tracy, Sandra Virtue, Yuhtsuen Tzeng & Paul van den Broek. 2004. Fluctuations in the availability of information during reading: Capturing cognitive processes using the landscape model. *Discourse Processes* 37(2). 165-186.

McCawley, James D. 1978. Conversational implicature and the lexicon. In Peter Cole (ed.), *Syntax and semantics, volume 9: Pragmatics*, 245-260. New York: Academic Press.

McCawley, James D. 1981. *Everything that linguists have always wanted to know about logic – but were ashamed to ask.* Chicago: University of Chicago Press.

McCloskey, Michael E. & Sam Glucksberg. 1978. Natural categories: Well defined or fuzzy sets? *Memory & Cognition* 6(4). 462-472.

Mercer, Neil McK. 1976. Frequency and availability in the encoding of spontaneous speech. *Language and Speech* 19. 129-143.

Mervis, Carolyn B. & Eleanor Rosch. 1981. Categorization of natural objects. *Annual Review of Psychology* 32. 89-115.

Mey, Jacob L. 2001. *Pragmatics: An introduction*, 2nd edn. Malden, MA: Blackwell Publishers.

Murray, Laura L. 2010. Distinguishing clinical depression from early Alzheimer's disease in elderly people: Can narrative analysis help? *Aphasiology* 24(6-8), 928-939.

Neale, Stephen. 1992. Paul Grice and the philosophy of language. *Linguistics and Philosophy* 15. 509-559.

Nieuwland, Mante S., Ditman, Tali & Kuperberg, Gina R. 2010. On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic

abilities. *Journal of Memory and Language* 63, 324-346.

Oxford English Dictionary Online. 2011. Oxford University Press. http://www.oed.com (September 2011).

Pustejovsky, James. 1991. The generative lexicon. *Computational Linguistics* 17(4). 409-441.

Pustejovsky, James. 1995. *Extensions to qualia structure.* Ms. Computer Science Department, Brandeis University.

Pustejovsky, James. 1998. *The generative lexicon.* Cambridge: MIT Press.

Récanati, François. 1991. The pragmatics of what is said. In Steven Davis (ed.), *Pragmatics: A reader*, 97-120. New York: Oxford University Press.

Richards, Meredith Martin. 1977. Ordering preferences for congruent and incongruent English adjectives in attributive and predicative contexts. *Journal of Verbal Learning and Verbal Behavior* 16, 489-503.

Rosch, Eleanor H. 1973. On the internal structure of perceptual and semantic categories. In Timothy E. Moore (ed.), *Cognitive development and the acquisition of language*, 111-144. New York: Academic Press.

Rosch, Eleanor H. 1975. The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance* 1(4). 303-322.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & B. B. Lloyd (eds.), *Cognition and categorization*, 27-48. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, Eleanor. 1981. Prototype classification and logical classification: The two systems. In E Scholnick (ed.), *New trends in cognitive representation: Challenges to Piaget's theory*, 73-86. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sayre, Kenneth. 1967. Philosophy and cybernetics. In Kenneth Sayre & Frederick J. Crosson (eds.), *Philosophy and cybernetics: Essays delivered to the philosophic institute for artificial intelligence at the University of Notre Dame*, 3-34. Notre Dame: University of Notre Dame Press.

Scott, Gary-John. 2002. Stacked adjectival modification and the structure of nominal phrases. In Guglielmo Cinque (ed.), *The cartography of syntactic structures*. Vol. 1, *Functional structure in the DP and IP*, 91–120. Oxford: Oxford University Press.

Shannon, Claude E. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 379-423, 623-656.

Smith, Carlota S. 1997. *The parameter of aspect*. Dordrecht: Kluwer Academic Publishers.

Smith, Edward E. & Douglas L. Medin. 1981. *Categories and concepts*. Cambridge, MA: Harvard University Press.

Sperber, Dan & Deirdre Wilson. 1995 (1st edn. 1986). *Relevance: Communication and cognition*, 2nd edn. Oxford, UK: Blackwell.

Stark, Jacqueline Ann. 2010. Content analysis of the fairy tale Cinderella — A longitudinal single-case study of narrative production: "From rags to riches." *Aphasiology* 24(6-8), 709-724.

Steele, Sara C. & Watkins, Ruth V. 2010. Learning word meanings during reading by children with language learning disability and typically-developing peers. *Clinical Linguistics & Phonetics* 24(7), 520–539.

Syrett, Kristen, Kennedy, Christopher & Lidz, Jeffrey. 2009. Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* 27, 1-35.

Verkuyl, H. J. 1972. *On the compositional nature of the aspects*. Dordrecht: D. Reidel Publishing Company.

Wang, Chin-An & Inhoff, Albrecht W. 2010. The influence of visual contrast and case changes on parafoveal preview benefits during reading. *The Quarterly Journal of Experimental Psychology* 63(4), 805-817.

Wasow, Thomas. 1977. Transformations and the lexicon. In Peter W. Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal syntax*, 327-360. New York: Academic Press, Inc.

Wilson, Deirdre & Dan Sperber. 2012. *Meaning and relevance*. Cambridge: Cambridge University Press.

Wittgenstein, Ludwig. 1953. *Philosophical investigations*. New York: Macmillan.

Wulff, Stefanie. 2003. A multifactorial corpus analysis of adjective order in English. International Journal of Corpus Linguistics 8:2, 245–282.

Young, John Frederick. 1971. *Information theory*. London: Butterworth.

Yule, George. 1996 (2008 printing). *Pragmatics* (Oxford Introductions to Language Study). Oxford, UK: Oxford University Press.

Zadeh, Lotfi. 1965. Fuzzy sets. *Information and Control* 8. 338-353.

Zipf, George Kingsley. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.