University of South Carolina
Scholar Commons

Theses and Dissertations

1-1-2013

The Relevance of Sarcasm In Resolving Ambiguous References In Spoken Discourse

Sara Ann Peters
University of South Carolina - Columbia

Follow this and additional works at: http://scholarcommons.sc.edu/etd

Recommended Citation

Peters, S. A.(2013). *The Relevance of Sarcasm In Resolving Ambiguous References In Spoken Discourse.* (Doctoral dissertation). Retrieved from http://scholarcommons.sc.edu/etd/2507

This Open Access Dissertation is brought to you for free and open access by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact SCHOLARC@mailbox.sc.edu.

THE RELEVANCE OF SARCASM IN RESOLVING AMBIGUOUS REFERENCES IN SPOKEN DISCOURSE

by

Sara Peters

Bachelor of Arts
University of South Carolina, 2006

Master of Arts
University of South Carolina, 2010

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Experimental Psychology

College of Arts and Sciences

University of South Carolina

2013

Accepted by:

Amit Almor, Major Professor

Anne Bezuidenhout, Committee Member

Fernanda Ferreira, Committee Member

Douglas Wedell, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

© Copyright by Sara Peters, 2013
All Rights Reserved.

DEDICATION

To Mom, for her emotional support; to Dad, for his emotional and financial.

To Amit, for years of patience.

ABSTRACT

Sarcasm, or sarcastic irony, involves expressing a message that is often opposite of the literal meaning of what is being said, in a way that may sound bitter, or caustic (Gibbs, 1986). In the past, sarcasm has been viewed as a method of introducing the possibility of alternative interpretations of a discourse, by creating ambiguity as to the intended discourse interpretation. The current series of experiments sought to demonstrate that sarcasm could be viewed as *beneficial* in resolving ambiguity in conversation, by highlighting particular interpretations and thus ease processing, dependent on other available contextual information. Two Visual World studies are reported in which this theory is tested. First, the variables associated with the social contexts represented in the conversations were normed in Experiment 1. Second, spoken conversations involving two speakers discussing events that were occurring within a town were presented to participants in Experiments 2 & 3. Experiment 2 presented a two-sentence conversation in which the first speaker introduced an ambiguous homophone in their utterance, and a second speaker followed with a comment made using Sarcastic Prosody. Experiment 3 also presented a two-sentence discourse, with the first speaker making a generic comment, and the second speaker following with a homophone reference spoken with Sarcastic Prosody. Within the experiments, sarcasm increased the processing of alternative interpretations of the homophones differently depending on the social context and the characteristics of the homophone (such as written frequency, and meaning dominance), suggesting it successfully highlighted particular alternatives, rather

than all possible interpretations. Theories such as Relevance Theory would predict this effect of sarcasm, such that given the proper conversational and contextual constraints, sarcasm can be used by speakers in a manner beneficial to listeners.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Spoken language messages offer information to listeners both through the composition of the message, and in additional speech, speaker, word and discourse-level information that may be relevant. Examples include the prosody used to impart the message (particularly the intonation of particular utterances), the emotion portrayed by the speaker, and the surrounding situational context for the information. Utilized successfully by a listener, this additional information may provide indirect information on a speaker's attitudes and beliefs about a current topic, as well as information regarding any underlying meaning of the message (Kreuz & Glucksberg, 1989; Capelli, Nakagawa, & Madden, 1990; Weingartner & Klin 2005; 2009). According to Grice's Cooperative Principle, conversation interlocutors choose what they say so as to further the purpose of the conversation, and we expect this principle to be understood by both parties as they engage in the act (Grice, 1975). This is the case even when an utterance from one of the parties may at first seem out of place with the preceding context, as is often the case with sarcasm, or sarcastic irony.

When defining sarcastic irony, Gibbs (1986) uses the definition offered in the *Oxford English Dictionary*, in which irony is "the use of words to express something other than and especially the opposite of the literal meaning of a sentence", and *sarcastic* irony is defined as "bitter, caustic, and other ironic language that is usually directed against an individual". These definitions provide a useful starting point in the

consideration of sarcasm as a linguistic device actively chosen by writers and speakers for use during conversation and text to serve a *specific* purpose in the discourse, to alter the interpretation further from the literal meaning toward an intended meaning. This suggests that when sarcasm is chosen by a speaker and used in a spoken language context, it can be geared toward providing information for the conversational partner, or other party in a specific situation. For example, if a word such as "bulb" was used in a conversation, which is a homophone with multiple meanings, sarcasm may be used to highlight the meaning of "flower bulb", rather than "light bulb". Thus, in a given context or conversation, if sarcasm is being used as a linguistic device in this manner, we should find that the use of sarcasm is highly dependent on the preceding context as well as the situation, and highlights a particular interpretation given these constraints. We should also expect that theories of sarcasm use should be able to account for this contextual (situational) sensitivity.

Multiple theories have been put forth to explain the use and resolution of irony in discourse and in general, with a few paying special attention to the use of sarcastic irony, or sarcasm (Kreuz & Glucksberg, 1989; Dews & Winner, 1995; Giora, 1995, 1997; Gibbs, 1984, 1989, 1994, from Gibbs 2002; Sperber & Wilson, 1986). The current project is unique in that it aimed to examine theories based on their ability to account for data related to one purpose that sarcasm can serve during a discourse, which is the use of sarcasm to *aid* in resolving an ambiguity created by a homophone in spoken discourse. Another way of describing this goal is that the current project aimed to demonstrate that sarcasm could be viewed as beneficial from the point of a speaker, if it highlights the correct meaning of the utterance for a listener amongst alternatives. Previous work has

shown that sarcasm itself can successfully introduce ambiguity into a discourse by creating possible alternative interpretations for listeners (Weingartner & Klin, 2009; Klin & Drumm, 2010). In this previous work, the sarcasm served to delay processing by introducing competing interpretations. However, work examining prosody changes in spoken language has found that listeners take stress information to indicate that the phrase or word is important, and that the stress information can alter lexical processing (Cutler, Dahan, & van Donselaar, 1997; Blutner & Sommer, 1988). Since one of the components that can indicate sarcasm during spoken language is the prosody (cues such as intensity changes, duration changes, etc. (Rockwell, 2002; Cheang & Pell, 2008)), if the cues are taken to indicate a stress on an important phrase, listeners may utilize this information and change their interpretation of a discourse. The current work predicts that this effect of sarcasm can be observed systematically, such that the increased processing will lead particular discourse interpretations to become prominent, and more likely to be considered.

Theories of sarcasm resolution currently in the field exist on a spectrum. Some work presents arguments for theories of a specific sarcasm resolution mechanism, while alternatives place sarcasm in the context of more general language processing (Dews & Winner, 1995; Kreuz & Glucksberg, 1989; Gibbs, 1986, Giora, 1997; Sperber & Wilson, 1986, 1995). Theories such as the Echoic Reminder Theory and the Tinge/Muting Hypothesis suggest that sarcasm either echoes a violated social norm, or mutes criticism or praise, respectively (Kreuz & Glucksberg, 1989; Dews & Winner, 1995).

Another theory of the processing of irony - and therefore sarcasm - the Direct Access View, suggests that the sarcastic interpretation of an utterance should evoke no

additional cognitive load for a listener (Gibbs, 1984; Gibbs, 1989). This view suggests that the processing of sarcastic versus sincere statements should have similar time frames, with irony simply a different way of transmitting the message (Gibbs, 1986). The Graded Salience Hypothesis (Giora, 1997) is an integrated language processing account, suggesting that the salience of a particular interpretation of an utterance is weighted when resolving whether the utterance is meant literally or not. In this approach, it is aspects of the message including the conventionality (Gibbs, 1980 from Giora, 1997), familiarity of the expression (Blasko & Connine, 1993 from Giora, 1997), and the frequency (Hogaboam & Perfetti, 1975; Neill, Hilliard, & Cooper, 1988, both from Giora, 1997) and status of the message within the context, that affect the salience of each meaning (literal versus nonliteral, in cases such as "The grass is greener on the other side."), and the final interpretation of the utterance.

A last approach placing sarcasm processing in a general language-processing framework, that is able to account for more than just sarcasm and other forms of irony is Relevance Theory (Sperber & Wilson, 1986, Sperber & Wilson, 1995). Relevance Theory expands upon Grice's previously mentioned Cooperative Principle, and outlines further assumptions regarding the expectations of both parties as they engage in speech (Sperber & Wilson, 1986; Sperber & Wilson, 1995). While some information may be provided by speakers in a straightforward manner, Relevance Theory postulates that all aspects of utterance chosen by speakers are typically meant to offer listeners cues relevant to the speaker's message, and yet still, that this is understood by both parties. This is the case even when a message at first appears ambiguous, or contrary to expectations, as is often the case with sarcastic irony. Relevance Theory (Sperber &

Wilson, 1986; Sperber & Wilson, 1995) supports the notion that the speaker deems any additional processing used to resolve a sarcastic utterance worth the effort when the utterance is chosen.

Alternatively, some work on the process of conversational adaptation by partners has shown that a speaker may not always be mindful of the interlocutor's view and needs during conversation. In these cases, it is hypothesized that the information maintenance of common ground expends too much effort (Pickering & Garrod, 2004). Thus, listeners may not always take advantage of the information offered by speakers. However, it's noted that the speaker does have some tools to highlight important information for interlocutors, such as accentuation and deaccentuation of salient information using prosody, as previously mentioned (see Cutler, Dahan, & van Donselaar, 1997 for review of work). Thus, using a Relevance Theory approach, it may still be possible that the contradictory effects of sarcasm use (the use of sarcasm by a speaker, even though it may appear to introduce ambiguity, and to violate the Cooperative Principle) could be justified if prosody cues are given to highlight the information, suggesting it is important. This assumption allows us to consider that sarcasm is chosen for a *specific purpose*, to perhaps identify a particular interpretation, or increase processing of an alternative.

Each of the approaches described above will be further explored in the current work. For a subset of the theories, namely Graded Salience and Relevance Theory, generalized predictions for outcomes in the current set of experiments will be generated and evaluated. For this subset, particular attention is paid to how well each approach accounts for the findings in the wider literature. The remainder of the study is outlined as follows. First, Chapter 2 provides information about the nature of sarcasm and our ability

to resolve sarcastic utterances, with a more in-depth exploration of the theories introduced above. The manipulation that the current work examines, the effect sarcasm contributes to ongoing discourse interpretations that necessitate resolving ambiguity generated by a homophone reference, is also introduced. Following, the methodology the current work utilized is discussed. Additionally, since sarcasm was explored as contributing information for successful ambiguity resolution, Chapter 3 describes a method to introduce ambiguity into the discourse, namely the use of homophones. Chapter 4 provides information on the final aspect of the chosen methodology, the Visual World Paradigm, with the remainder of the work focused on reporting the specific contributions the experimental sequence offers to the current literature.

CHAPTER 2

SARCASM

As previously mentioned, sarcasm is typically defined as "bitter, caustic, and other ironic language directed at an individual" (Gibbs, 1986). The approach taken in the current work is that sarcasm is a linguistic device chosen by speakers for specific purposes determined by the context of the conversation or discourse. It is noted that contexts vary greatly, with variables such as different social settings (situational contexts), and the frequency of the words used to describe the context, etc. contributing to how effective sarcasm will be as a manipulation. Additionally, here sarcasm is operationalized by the use in an utterance containing prosody cues previously identified by native English speakers as indicating sarcasm (Rockwell, 2002). It follows that if speakers have access to the use of sarcasm, in order for it to be an efficient choice for communication, we must also have the ability to rapidly resolve sarcastic utterances and utilize the information they offer during the discourse interpretation process. In order to support this assumption, the ability of native speakers to detect sarcasm must be verified, and questions regarding individual differences noted.

2.1 BEING NORMAL

Readers and listeners can describe specific characteristics of sarcasm that allow for identification and detection. In particular, previous work in the spoken language domain has noted that both children and adults identify sarcasm using prosody and context, although perhaps to differing extents (Rockwell, 2000; Cheang & Pell, 2008;

Bryant & Fox Tree, 2002). For example, when children are asked to explain how they detected sarcasm during a spoken discourse, they may report that the speaker "sounded mean" (Capelli, Nakagawa, & Madden, 1990; Nakassis & Snedeker, 2002). In addition, sarcasm is apparently more difficult for younger children than older children, and the ability to detect it properly thus seems to develop with age (Sullivan, Winner, & Hopfield, 1995).

In terms of prosody in particular, work by Rockwell (2000) identifies that sarcasm is typically indicated by a slower *tempo* (speech rate), spoken with greater *intensity* (amplitude), and at lower *pitch* level. More recently, work by Cheang & Pell (2008) affirmed these cues, noting that in their work, intensity appeared to be the most important cue. Work by Peters, Wilson & Almor (2012) has also confirmed native speakers' ability to detect sarcasm using both prosodic and contextual cues, within the available South Carolina participant pool that the current work utilized. The prosody cues were manipulated in sentences that were recorded in sincere tones, and the same procedure was used in the current experiment to achieve this effect with the cues noted above.

Thus, previous research shows that normal, native adult English speakers appear to have gained the knowledge necessary to aid them in resolving the true meaning of a sarcastic utterance. In particular, native English speakers will use both the context and prosody information present when the utterance is presented as spoken language (Bryant & Fox Tree, 2002). Additional work (beyond the scope of the current study) suggests that this knowledge may exist in the form of abilities that individuals obtain on a normal developmental trajectory, including the ability to take another's perspective and to recognize other's emotions (Happé, 1993; Sullivan, Winner, & Hopfield, 1995;

Chevallier et al., 2011). Because both context and prosody are used by many types of comprehension processes besides the detection of sarcasm (Cutler, Dahan, & van Donselaar, 1997; Wilson & Wharton, 2006), the conclusion of this brief review of "normality" suggests that theories of sarcasm use should strive to place it in the context of a language mechanism that relies on general (rather than specific) processing abilities to resolve.

2.2 Theories Of Sarcasm Use

Theories of sarcasm use and resolution are situated in several different areas in the existing literature, including the study of social norms and their violations, work on Theory of Mind, and work on non-literal language processing. Each of these areas emphasizes different mechanisms and reasons for the use of sarcasm in discourse, and is impacted when there are impairments in the ability to detect sarcasm (or irony) (Happé, 1993; Chevallier et al., 2011; Dews & Winner, 1995; Kreuz & Glucksberg, 1989; Gibbs, 1986; Giora, 1995, 1997; Sperber & Wilson, 1986). Sarcasm is often found to refer to social norms, current situational processes, and the actions of others within a negative (or positive) context, and failure to detect it has practical implications both in specific conversational settings, and for the development of language processing abilities. Research in each of these areas is generally focused on the issues central to the area, thus leaving the role of general linguistic mechanisms largely unaddressed. For the purpose of the current work, sarcasm as a linguistic mechanism is further defined as a mechanism that has the ability to *enhance* discourse processing by offering particular alternative interpretations, given that participants recognize the relevant information (such as attitudes, beliefs, violated norms) successful resolution offers, and speakers intend them

to do so. Therefore, any processing costs of sarcasm are mediated by the addition of this additional information, which may or may not be activated for a listener or reader in the same timeframe as a (series of) sincere statement(s). Although the focus here is on theories of resolution, these theories may also be considered from the speaker's initial viewpoint: that successful resolution is offering additional information, which is why sarcasm is chosen. This definition will aid in clarifying the placement of the current work using existing theoretical accounts of sarcasm use.

Before examining these approaches more thoroughly, several main assumptions underlying the current work are explicitly stated, which have been previously hinted. The first is that (1) sarcasm is one of many available choices of mechanisms speakers may utilize when preparing and executing an utterance. Speakers can theoretically convey their message either through sarcasm or literal, sincere statements, and have a similar outcome in terms of the information transmitted by utterances. Therefore, sarcasm, or sarcastic irony, must be chosen by speakers for a particular reason. The second assumption is that (2) there need not be dedicated mechanisms for explaining its use, rather that theories of sarcasm use and resolution should fit in broader language processing theories. Therefore, this work postulates that sarcasm is a device chosen by speakers because it serves the discourse function with the *best return of effort for the current context*. This assumption will also aid in clarifying the placement of sarcasm within the upcoming theoretical accounts, such that it should be beneficial in some cases, and irrelevant in others. This is not, however, a new approach, and is clearly not the only approach. Several theories of sarcasm use have already examined different reasons why sarcasm may be chosen by speakers, in terms of what may be gained through successful

resolution of the utterance (examples include: Tinge/Muting Hypothesis, Echoic Reminder Theory, Mention Theory of Irony (Dews & Winner, 1995; Pexman & Olineck, 2002Kreuz & Glucksberg, 1989; Jorgensen, Miller, & Sperber, 1984). Finally, once again (3) sarcasm is operationally defined as being identifiable by the prosody cues, particularly those previously noted by Rockwell (2000) & more recently Cheang & Pell (2008), and also through social contextual effects, which will be further outlined in section 6.10.

Echoic Reminder Theory (Kreuz & Glucksberg, 1989) broadly suggests that speakers choose sarcasm to remind an interlocutor that a social norm is being violated. In this way, sarcasm is meant to echo the social norm being violated, and is phrased as referring to such. For example, in a situation in which one friend (Person B) is late meeting another (Person A) at the movies, utterance [1] delivered in a Sarcastic Prosody is meant to convey to the interlocutor of Person A that norm [2] has been violated by Person B, and that Person A is not appreciative of that fact. In order to resolve this message, A's interlocutor must realize that [1] is meant as sarcasm by placing the sarcasm in the context offered against a background of social norms.

[1] He sure is punctual.

[2] It's not okay to be late to a meeting.

In this way, Echoic Reminder Theory begins to address the social aspect of the use of sarcasm, introduced in the discussion of context above. It requires the assumption that there are social contexts in which sarcasm may be perceived as more appropriate. These contexts could include variable subjects of discussion in a conversation, and not just social norms (for example, individuals may be more likely to make a sarcastic

statement regarding someone else's clothing, or hairstyle). Here, the Echoic Reminder Theory becomes less useful, particularly when sarcastic utterances move beyond the scope of reflecting social norms, and are reduced to simply a speaker's personal judgment of another, or individual differences of humor style. For example, previously sarcasm was described as being in part "bitter...directed against an individual" (from Gibbs, 1986). This definition does not suggest there always need be a social norm against which the judgments of the individual are tested. If a speaker does not like a coworker, he might refer to the individual's work as "just amazing" using sarcasm to an interlocutor, and the work may actually *be* amazing by conventional standards. Here, in order to understand the speaker's sarcastic intent, the listener would need to realize that the speaker does not like the subject of the utterance, which is the information leading to the sarcastic interpretation. So while the Echoic Reminder Theory (Kreuz & Glucksberg, 1989) may be useful in specific contexts, it does not appear to be able to handle the *full* range of social contexts in which sarcasm can be successfully employed. In order to do so, we would have to consider that the statements being echoed could be *very* distant and broad.

A second echoic theory, the Mention Theory of Irony (Jorgensen, Miller, & Sperber, 1984) also appeals to echoes of statements or sentiments made by a previous speaker as a method of intending an ironic, or in this case sarcastic, utterance. Jorgenson and colleagues (1984) suggest that the echoes themselves could seem remote from the statements that triggered them, and thus not necessarily appear as a straightforward echo of a previous utterance (e.g., they could include echoes of emotion, etc. leaving further room for interpretation). Mention Theory has utterances resolved by comprehenders

locating the expression the speaker's echo referred to, and then determining the speaker's attitude toward the topic (Jorgensen, Miller, & Sperber, 1984; Bryant & Fox Tree, 2002). From this perspective, [1] may appear remote because [2], while a known social rule, may not have been mentioned during conversation at any point that night. Nonetheless, the sentiment of [2] would likely be recognized and [1] resolved. However, this too involves knowledge of what we could refer to as "social" contextual knowledge norms in conversation, although not the "social norms" of the Echoic Reminder Theory.

Another theory that could explain [1] is the Tinge, or Muting Hypothesis (Dews & Winner, 1995; Pexman & Olineck, 2002). Using the Muting Hypothesis, sarcasm mutes what would otherwise come across from a speaker as outright negativity (or in other words, "tinges" it with the opposite emotion, positivity in this case). In addition, the Muting Hypothesis also suggests that when a compliment is given, the positive tone of the compliment is also muted. Therefore, sarcasm serves to mute the emotional aspect of either censure or praise. In this case, instead of indicating a direct reference to [2], Person A is muting [3] when conveying their feelings to their interlocutor.

[3] Being late is extremely rude, and I am upset with Person B.

Once again, while the Muting Hypothesis may apply in some contexts, it is possible to conceive of a situation in which, for example, a boss uses sarcasm to censure an employee in front of a group. When social dynamics are entered into the resolution of the message, the group may perceive that the censure is actually harsher than an outright censure that is not veiled in meaning, introducing the emotion of shame into the subject of the utterance. Social dynamics are mentioned repeatedly here, as they provide cues for resolving sarcasm that have not been studied as extensively (i.e., here we know that the

boss was being sarcastic by viewing the subject of the utterance's response to the comment). Thus, we are left with a theory that may provide useful information about sarcasm resolution in a specific context, but does not have the flexibility needed to account for the wide range of situations in which sarcasm can hypothetically be used.

A theory considered in more depth in the predictions of the experiments offered in this study, is Direct Access View (Gibbs, 1986; Gibbs & O'Brien, 1991, Gibbs, 2002). The Direct Access View offers information on the time course of the resolution of what would appear to be non-literal, or figurative interpretations of an utterance. Oftentimes, a sarcastic utterance can be detected because it is the opposite of what a speaker actually feels. In this situation, the opposite meaning of the phrase needs to be accessed and understood to interpret and resolve the utterance as the speaker originally intended. The Standard Pragmatic view of processing these types of messages suggests that listeners first interpret the literal statement before converting it to a non-literal interpretation (Grice 1989, from Gibbs, 2002). The Direct Access View (Gibbs, 1984, Gibbs, 1989), however, claims that there are situations in which listeners and readers can understand the figurative interpretations directly, and do not require time to reevaluate the utterance (Gibbs, 2002). Thus, the interlocutor of Person A may be able to identify [1] as sarcasm immediately, particularly if they have been waiting on Person B for five or ten minutes. Thus, Direct Access View does begin to address the social contextual information that appeared to be lacking in the previous approaches.

The previous explanations discussed view sarcasm as a choice made by speakers to convey their attitude to an interlocutor. In contrast, the Indirect Negation View of Irony (Giora, 1995; Giora, 1997) emphasizes the role of the listener in addition to the

speaker, and also suggests processing cost as a factor involved in resolution of utterances. This allows the Indirect Negation View to take into account the social dynamics involved in the production and comprehension of sarcasm, without requiring additional processing mechanisms.

The Indirect Negation View of Irony is different from the Direct Access View previously discussed, which suggests the processing of an ironic statement may take no longer than that of a literal statement (Gibbs, 2002; Gibbs & O'Brien, 1991). The Indirect Negation view asserts that both the literal and intended meaning in an ironic statement are activated, and that statement is resolved by processing the *difference* between the two meanings, and that this involves an additional processing load. Giora (1995) views direct negations as offering *different information* than indirect negation; further suggesting that indirect negation may not refer to the opposite of the literal meaning of a statement, as direct negation would. Returning to the example of a boss censuring an employee in front of others, if he (as the speaker) comments "That was some great work," using sarcasm may indicate more than just it was poor work, but instead that it was *horrible* work, the worst he has ever seen. Thus, Indirect Negation can account for a context that the Muting Hypothesis had difficulty explaining. This view of Indirect Negation is supported when considering that in sarcasm processing, additional information such as the emotion or attitude of the speaker is conveyed with the utterance when it is resolved as the speaker intended, and this information is not simply the opposite of the literal statement.

Giora's (1995) approach is based on situations in which her Conditions for Discourse, Irony, and (when it applies) Joke Well-Formedness are met. In terms of Irony

Well-Formedness, the ironic statement must: (1) Introduce information about accessible discourse topics, (2) Violate graded informativeness, and introduce a less probable message than the one expected in the given context, and (3) Make the addressee evoke an implicature between the marked and unmarked messages, whereby they can note the difference between them and conclude that the statement is ironic (Giora, 1995). The Indirect Negation View of Irony also suggests that irony may be used in situations where norm violations of all types are apparent (not limited to social norms). Thus, this theory has more explanatory power than the Echoic Reminder Theory (which pertains to social norms), and can account for data supporting both the Mention Theory of Irony (where echoes can be remote) and the Tinge/Muting Hypothesis (by situating the experimental items such that the contexts fall in situations benefiting from a Tinge/Muting of praise or criticism). Thus, it is general enough to be applied to multiple contexts, yet specific enough to provide predictions and reasoning for sarcasm usage by a speaker.

An evolution of sorts of Giora's Indirect Negation View of Irony, the Graded Salience Hypothesis (Giora, 1997), suggests that the *salient* meanings of figurative language statements are processed first, regardless if it is the literal interpretation or the figurative one, and that additional meanings are only processed after this is completed. In terms of sarcasm, this view suggests that [1] may first be processed as indicating that the person is punctual, before it is reevaluated to the sarcastic meaning that is meant. Additionally, other information already considered a salience cue (such as written frequency of a word) cannot be influenced or bypassed in processing by a context (Giora, 2002). Additionally, the order of processing is dependent on the other information provided by preceding discourse. If there is enough preceding information to indicate

that the sarcastic meaning should be taken as having more salience, it may be processed first. The information to direct this decision may come in the form of a variety of cues offered by the context and the statement itself. As the Graded Salience view is an extension of the Indirect Negation approach, the work here will only evaluate the Graded Salience view.

Finally, a last alternative, the aforementioned Relevance Theory (Sperber & Wilson, 1986; Sperber & Wilson, 1995), can account for sarcasm and also offer predictions for additional aspects of conversation and language processing. This makes Relevance Theory at least as flexible as the Graded Salience Hypothesis in terms of the predictions it can offer regarding the contribution of sarcasm to discourse processing. Again, the origin of Relevance Theory can be traced back to Grices' Maxim of Relevance, which states that speakers chose their utterances so as they remain relevant to the listener and the conversation (Grice, 1975).

Relevance Theory establishes conditions for the relevance of new information, which can be summed as having a purpose during the discourse (Sperber & Wilson, 1986; Sperber & Wilson, 1995). One of the most "essential" conditions for relevance as defined by this theory is what Sperber and Wilson (1995) refer to as "contextual effects". A contextual effect is based on the concept of what is considered a modification to the context of a situation or conversation. Sperber and Wilson (1995) refer to information that does not completely overlap with previously known information, yet is also not completely unrelated to the previous context, as the types of information likely to be able to modify a context enough to allow contextual effects. When new information meets these criteria, and removes an assumption about the context or situation that was

previously held, (and thus also weakens the linked assumptions) a contextual effect has occurred (Sperber & Wilson, 1986; Sperber & Wilson, 1995). While multiple types of contextual effects have been suggested, the use of sarcasm would appear to fall under two main categories: 1) Strengthening of previous information, or 2) Contradiction of previously held assumptions (Sperber & Wilson, 1986; Sperber & Wilson, 1995). It is likely that other factors of the context and information determine when sarcasm can serve these contextual effects.

Broadly, current versions of Relevance Theory suggest instances of sarcasm used during a discourse are expected to serve a purpose, or more loosely, be seen as "relevant" only given certain contextual constraints. This follows from the assumption that people have intuition regarding relevant versus irrelevant information related to the context (Sperber & Wilson, 1986; Sperber & Wilson, 1995). The finding that marked prosody is interpreted as information that may be important to the discourse (once again, see Cutler, Dahan, & van Donselaar, 1997 for review of work in this area), may serve as a mechanism for explaining this apparent "intuition".

While research on the time to resolve non-literal, or figurative utterances is controversial, Relevance Theory provides a method of explaining the apparent contradictions within the literature (Gibbs, 1986; Gibbs, 2002). The return on cognitive effort investment issue is addressed in Relevance Theory by the notion of *degrees* of relevance. When determining the relevance of information, multiple factors are taken into consideration, including the benefit of the information (in terms of the contextual effects on the previously held assumptions) and the mental effort expanded. In the case that processing effort is large and all other aspects of the context are equal, the situation

with greater processing cost will in turn yield a smaller value for the relevance of the information (Sperber & Wilson, 1986; Sperber & Wilson, 1995). We see then, that the intuitive balance that users of sarcasm may seek to strike between cost and benefit is supported in this framework using the "amount" of relevance sarcasm has (Sperber & Wilson, 1986; Sperber & Wilson, 1995).

Returning to the example introduced above, under Relevance Theory, either [2] or [3] may have motivated Person A to say [1], but as long as Person A's interlocutor resolves that [1] is sarcasm, the conversation can continue with the intended meaning of the speaker identified by the listener; interlocutor will realize that Person A is upset with Person B. The longer Person A has waited with his interlocutor before issuing utterance [1], the more generally "relevant" the information may become. The usefulness of this approach extends to metaphor or to ambiguity resolution, insofar as if the ambiguity is resolved, productive conversation continues, although this is not to dismiss the possibility of conversation continuing without a correct resolution. However, if resolution is incorrect, the possibility exists that the listener's discourse interpretation may remain functionally incorrect, although it may not hinder the conversation from being "carried out".

The theories reviewed above all describe sarcasm resolution in the greater context of discourse or conversation processing. These theories, to differing extents, suggest that the resolution of the sarcasm is intended to have effects on the expectations of readers or interlocutors, whether it be in how they view a particular action or viewpoint (e.g., "He did some *great* work."), or character (e.g., "Joe is such a *hard worker*.") (Sarcastic Prosody indicated by italics). However, given the way that previous work that has

utilized sarcasm in experimental designs, less is known about how the resolution of the sarcasm itself affects interpretations of the overall discourse or conversation beyond the utterance. For example, there is a question of whether the use of sarcasm changes a listener's or comprehender's more global interpretation of a discourse. In other words, is the use of sarcasm during a discourse powerful enough to overcome other discourse information? Examples of this information include possible inherent bias in the conversation, with which the topics discussed are more reliably linked to a particular interpretation (e.g., perhaps "Joe" from above, is really a clown around the office, but still a hard worker.), or the frequency with which the topics are encountered in everyday language differs such that one interpretation is preferred. To examine this, we can briefly look to discourse functions of sarcasm that *have* been studied, to determine the extent to which they have thus far informed overall knowledge regarding the resolution process.

2.3 Discourse Functions Of Sarcasm

While some discourse functions of sarcasm have been identified, and are used during studies as discourse manipulations (Weingartner & Klin, 2009; Klin & Drumm, 2010), the limits of these functions are not well understood. That is, the focus of these studies was not the discourse function of sarcasm per se, but rather what these manipulations offer in terms of creating situations in which participants or listeners must draw inferences and give responses based on those inferences. Another way of considering this is to state that sarcasm has previously been used to *create*, rather than *resolve*, ambiguity by being responsible for introducing alternative discourse interpretations. (Weingartner & Klin, 2009; Klin & Drumm, 2010). Within this creation, there were no *a priori* expectations of the contribution of sarcasm to resolution patterns;

which is the approach taken here. Instead, it was assumed that differences would be created in situations where sarcasm was used because alternative interpretations would be considered, versus situations where it was not.

Weingartner & Klin (2009) used sarcastic statements within a text-based study involving discourse processing to determine whether or not readers would interpret sarcastic statements based on the knowledge available to them about the character's feelings (as the reader), or the knowledge available to other characters in the story (which was limited). In this way, sarcastic statements were used to introduce ambiguity into a character's statement, and observe whether readers interpreted it as such, or deferred to previous information. Hypothesizing why a speaker would choose to employ sarcasm to create alternative explanations, or introduce ambiguity, is beyond the scope of the current work, but the use of sarcasm by Weingartner & Klin (2009) serves as a another demonstration of its flexibility as a mechanism that can aid in our understanding of normal language processing.

In the studies reviewed above, although sarcasm had a functional purpose in the experimental designs, it is difficult to determine the impact sarcasm alone had on the readers' resolution of the characters' statements (Weingartner & Klin, 2009). Some of the contexts were likely more socially favorable to the use of sarcasm. In addition, the study looked at how information was attributed, whether or not readers understood that a given character might not have access to another's motivations. The goal was to create differences in interpretation, not explain why they were occurring.

Basic research, regarding how sarcasm affects discourse interpretations in the presence of other controlled information, is important in determining how strong of an

effect sarcasm presents as a discourse manipulation, or informational component in a discourse or utterance, should an author or speaker choose this method of delivery. Essentially, additional control is needed over the discourses to determine how sarcasm influences resolution processes: information regarding both the sarcastic statement itself, and in turn the discourse that it is a part of. Another way of phrasing this, is noting that we need to examine how sarcastic statements lead to the resolution of ambiguity within discourse and conversation; speakers and writers (often) mean for sarcasm to lead to a fixed interpretation when utilized. Thus, within the design employed, which includes homophones to create ambiguity and possible alternative interpretations, we should expect certain differences of the effect of sarcasm given characteristics of the homophone itself, and the context it is placed within. This question then also allows the application of the previously reviewed theories of sarcasm use and resolution to be applied to the results.

Individual Differences?

In terms of sarcasm processing, individual differences have not been extensively studied beyond work looking to identify gender differences in irony usage (Colston & Lee, 2004). Data reported in a second study was correlational and self-reported, pertaining to usage and humor style (Ivanko, Pexman, & Olineck, 2004), and both studies had effects that were not considered stable. However, the impact of (at this point) hypothesized individual differences in sarcasm processing, has been noted in the sarcasm literature. One example is areas that aim to use the differences to explain differences in normal versus abnormal processing of language in special populations (Happé, 1993; Sullivan, Winner, & Hopfield, 1995; Chevallier et al., 2011; Colston & Lee, 2004).

Since the current work provided an opportunity to collect data on some aspects of individual differences, data from a task measuring social contextual awareness was collected. This is important when considering sarcasm processing as a "normal" ability. Gender information was collected as well, but due to the high female to male ratio within the data, the effects were not analyzed.

Summary

The current project aims to look at sarcasm in situations where it was intended as *beneficial* in conversation; an added piece of useful information when resolving the ambiguity created by the use of a homophone. That sarcasm can ease comprehension is compatible with Relevance Theory, which emphasizes the function of various forms of language, not just their processing cost (Sperber & Wilson, 1995). It is also compatible with Graded Salience, which suggests that information that can affect salience is likely to be considered during ongoing discourse processing (Giora, 1997). As such, these two theories will receive the majority of the focus within the current work. As previously stated, this function of sarcasm that has not been studied extensively: the use of sarcasm as a method contributing to *resolution* of ambiguity in a discourse, by activating a particular interpretation, or increasing the salience of a particular alternative interpretation. By examining this question using a paradigm involving spoken language, we can clarify the constraints that create a context in which sarcasm successfully highlights particular alternatives. This investigation aids in discerning the full impact of sarcasm during discourse processing, and offers additional information on the wide range of reasons why speakers may specifically, and consciously choose to employ it.

CHAPTER 3

AN INTRODUCTION TO AMBIGUOUS WORDS & HOMOPHONES

In order to examine sarcasm as contributing to ambiguity processing, or more specifically homophone processing, several points need to be clarified in terms of what is meant by "ambiguity" within this work. Here, there is considered to be "ambiguity" within a discourse when the possibility of multiple working interpretations of a discourse is present. Ambiguous words, or words that can be interpreted as having multiple meanings, have a variety of properties that make them ideal for introducing this type of an ambiguity in the current work.

Ambiguous words such as "bank" can refer to multiple meanings of the same word (e.g., either a financial bank or a river bank). Additionally, when a homophone (one of a pair of words that share a phonological representation, but may differ in spelling, or orthography) is spoken, different words such as "fir" and "fur" can be interpreted as referring to the intended or opposing meanings, depending on the context provided (Kerswell et al., 2007). Thus, when spoken, both of these types of words can create an ambiguity in how a discourse should be interpreted, given that no additional context is provided. The current design utilized a spoken language methodology to control manipulation of both the prosodic and contextual cues that individuals have reported lead to a sarcastic interpretation of a statement (Cheang & Pell, 2008; Rockwell, 2002; Bryant & Fox Tree, 2002). Since the stimuli were spoken, both ambiguous words that share phonological codes (where spelling does not differ), as well as homophones,

were used to generate an ambiguity resulting in multiple possible discourse interpretations (Kerswell et al., 2007). As both versions of ambiguous stimuli described previously fall under "homophone" in terms of sharing phonological codes corresponding to multiple meanings, the word "homophone" will be used to refer to both groups.

The use of homophones provided a context in which there were multiple possible interpretations of an utterance in a conversation, yet given the different characteristics of the homophone itself, some were more likely than others. Before the effect of sarcasm on resolving a homophone in a discourse can be evaluated, the likelihood of one interpretation of a homophone (and therefore a given discourse) (e.g., light bulb) versus another (e.g., flower bulb) must be understood. While measures of factors termed here as Dominance (which homophone meaning is more likely to be ascribed with no prior context) and Frequency (written corpus Frequency) are available for homophones, the process of how these multiple lexical and discourse level factors interact to resolve a homophone independently, as opposed to when sarcasm is utilized within a discourse, has not been studied. A baseline needs to be established as to the likely interpretation of a homophone given the interaction of the Dominance of one meaning of the homophone, the Frequency of the critical homophone in the statement, as well as the Social Context (an additional factor of the conversation), before the additional effect of sarcasm on this processing can be analyzed. Accounting for these characteristics of the homophones, and thus the possible interpretations of a discourse, allows us this baseline for comparison when Sarcastic Prosody is added to an utterance. Further introduction of each of the factors of importance of the homophone is given below.

Dominance

Homophones can have corresponding representations that are either Balanced in the amount with which the multiple meanings are used within language (e.g., "pitcher" referring to a baseball player or device to pour drinks, both occur relatively as often), or Biased, with one representation and corresponding meaning (or more meanings) encountered and suggested in norming studies more often than another (e.g., "ball", where a round toy is more common than a dance), (Swinney, 1979, Rayner & Duffy, 1986, & MacDonald et al., 1994; all from Mason & Just, 2007).

To clarify, if a homophone is Biased, we refer to one meaning of the homophone as Dominant (more likely) and the other, less likely, as the Subordinate. Previous work has used a familiarity rating process to establish dominance in sets of homophones (Kreuz, 1987; Morris (unpublished norms), 1995 - ongoing) within the populations of interest to the current work (college students, and the University of South Carolina participant pool). Here, measures of Dominance were combined from the two corpora to generate them for experimental items, as a full set did not exist from either source. Dominance of a meaning typically creates an initially higher level of salience for one meaning (given an unbiased context), that can either be affirmed with continuing cues (such as with a visual representation of the meaning), or not. Given that cues reinforcing more than one meaning, (e.g., pictures, or contextual representations of both the Dominant and Subordinate meaning in biased homophones) are present within the current work, it is expected that there will be differences in the effect of Dominance on discourse processing within the current experimental paradigm compared to previous work. It is also expected that Dominance will interact with a second cue present in homophones,

Frequency effects (explained below), when discourse interpretations are generated. For this reason, homophones that are Biased (and therefore contain initial Dominance differences) will be analyzed separately for effects than those that are Balanced (where meanings are equally available initially). An example of this division of homophones into categories can be viewed in Table 3.1.

Table 3.1 Sample homophones divided by Dominance and Frequency.

Homophone	Meaning	Dominance	Initial Salience	Frequency
Ball	Round Ball	Biased	Dominant	High
Ball	Dance	Biased	Subordinate	High
Sale	Event	Balanced	Equal	High
Sail	For Boat	Balanced	Equal	Low
Bulb	Light	Biased	Dominant	Low
Bulb	Flower	Biased	Subordinate	Low

Frequency

Frequency of usage in language of a particular word can be measured by the number of appearances of the word within a corpus. For example, just because when the word "jeans" is pronounced aloud, the Dominant meaning is blue jeans, and the Subordinate is genes as in genetics, this does not mean that either is common in general language usage. In the current study, Frequency is measured by frequency of occurrence in the Kucera & Francis corpus (1967). Effects based on the Frequency with which a term appears in language have been widely reported in the literature pertaining to both written and spoken language. (Cutler, Dahan, & van Donselaar, 1997; Dahan, Tanenhaus & Chambers, 2002; Rayner & Duffy, 1986) One reported effect is that low frequency

words are typically fixed on longer in reading paradigms (Rayner & Duffy, 1986). The addition of homophones which do not share corresponding orthographic representations offers the ability to estimate the Frequency of use of both words appearing in a standard corpus while being sure of their intended meanings (e.g., jeans vs. genes), which is impossible with strictly ambiguous words (e.g., ball). In the second case, Dominance information is combined with the Frequency categorization of the ambiguous word in a regressive design across items to better estimate the frequency of meanings (Griffin, 1999; Lucas, 1987). While Frequency estimates are corpus based written estimates, Dominance effects are thought to contain a listener's judgment of both frequency *and* appropriateness. Thus, while the effects are separated within the current design, the regressive nature allows the utilization of Balanced and Biased homophones, with written frequencies that are high or low (see Frequency category of Table 3.1). This process is further detailed in Chapters 6 & 7.

Other Information on Homophone Processing

While these two aspects of a homophone are expected to heavily impact processing within a discourse, there are additional effects that need to be considered as well. Evidence from behavioral studies has indicated that in terms of processing time, it takes readers longer to read sentences with an ambiguous word than those without (Duffy et al., 1988, Miyake et al., 1994, Rayner & Duffy, 1986; all from Mason & Just, 2007). This evidence suggests that during a reading paradigm, individuals are sensitive to polysemous homophones as they are encountered, and resolving the meanings delays the time course of processing. Additional work on homophone processing has noted that when they are used in spoken language, prosody interacts with their processing. In this

work, multiple meanings were activated if the words are placed in focus using accent patterns; also, the focused parts receive a more detailed semantic processing of their lexical meaning (Blunter & Sommer, 1988, from Cutler, Dahan, & van Donselaar, 1997). Thus, the search for the meaning of the full word and sentence is related to the interaction of the prosody with the information that becomes available during with lexical access of an ambiguous word.

In further work using *spoken* sentences combined with visual representations, participants have also shown sensitivity to the activation of multiple meanings of a homophone (Mirman et al., 2008b; Chen & Boland, 2008). Visual World Paradigm (hereafter VWP) experiments, which combine spoken language stimuli with a visual display, have been used to demonstrate the effect. Participants have been found to fixate on pictures corresponding to both meanings of an ambiguous word in nonrestrictive contexts (the effects are brief), and fixate more on one meaning when the context was restricted appropriately (Mirman et al., 2008b). In addition, both frequency and context effects have been noted to have a role in homophone resolution in the VWP as well (Chen & Boland, 2008).

The Current Work

Utilizing previous work, within the current study, the effect of sarcasm on resolving a homophone reference assumes an effect on the combined result of the Dominance and Frequency interaction present within each homophone, as the two are difficult to disentangle. The other effects of mention on homophone processing must be considered as well, such as the increased lexical processing (Blunter & Sommer, 1988, from Cutler, Dahan, & van Donselaar, 1997; Mirman et al., 2008b; Chen & Boland,

2008). However, when this information is considered, the framework of the Graded Salience (Giora, 1997) and Relevance Theory perspectives (Sperber & Wilson, 1995) presented in Chapter 2 can then be used to generate hypotheses which favor different homophone interpretations, when sarcasm is present or absent in an utterance.

If the effect of sarcasm interacts with the Dominance and Frequency of a particular homophone within a discourse that offers no additional contextual information, sarcasm may be preferred in one level the interaction (highlighting it), and it may have a negative impact in another, resulting in a processing delay that mimics those noted in other figurative language research (Giora, 1995; Giora, 1997). In this case, to evaluate differences in a Graded Salience versus Relevance Theory approach to the effect of sarcasm on discourse interpretation, we can examine whether sarcasm affects interpretation by slowly building salience towards a particular interpretation of the discourse (through additional processing of the meaning), or highlights a particular interaction of information, and thus rapidly changes the interpretation of the discourse (Giora, 1997, Sperber & Wilson, 1986; Sperber & Wilson, 1995).

One way of examining the interactions of these multiple variables is by using an eye-movement study paradigm, such the VWP. This paradigm allows the visual depiction of the different contexts of the homophones, and the online tracking of the processing of competing interpretations by examining attention as measured by eye-movements to the depictions. If Dominance and Frequency affect the interpretation of a homophone early on and persist, it's possible that sarcasm will have no effect on changing a listener's interpretation of a discourse, if it does not occur with the homophone. However, if at the onset of Sarcastic Prosody following a homophone

spoken sincerely, there are changes in eye-movement patterns, an argument can be made that the sarcasm is able to affect discourse processing, and perhaps interpretation as well. Exactly these types of data can be provided using the VWP, as described next.

CHAPTER 4

THE VISUAL WORLD PARADIGM

The Visual World Paradigm allows for the study of spoken language comprehension using an online method in which stimuli are spoken, and eye-movements to items related (or not) in the visual display are tracked (Cooper, 1974; Tanenhaus et al, 1995). Three important aspects of this paradigm: the discourse, visual display, and a linking hypothesis that connects the two for the current study, (Tanenhaus et al., 2000) can be manipulated adequately for use in the current study.

Because sarcasm is identified by adults using both context and prosody information, a task utilizing spoken language processing provide a valid area in which to study how well sarcasm functions to contribute to ambiguity resolution because both pieces of information can be utilized (Capelli, Nakagawa, & Madden, 1990; Nakassis & Snedeker, 2002; Rockwell, 2000). Spoken language can be studied using the VWP, which allows prosody and visual context to interact, and offers an eye-movement record as a measure of the online time course of processing. In addition, the ambiguity, frequency, and dominance effects introduced in the previous chapter have all been established to exist within the VWP (Mirman et al., 2008b; Chen & Boland, 2008).

In terms of discourses used in VWP studies, the paradigm has adapted very well to the testing of multiple processing components. Many versions of this paradigm have been used since it was first created, with the discourse serving a variety of functions. It has been used to compare what type of lexical information is accessed first upon hearing

a word (Allopenna, Magnuson, & Tanenhaus, 1998), when semantic information is accessed (Cooper, 1974; Huettig & Altmann, 2005; Huettig et al., 2006; Yee & Sedivy, 2006; Yee, Overton, & Thompson-Schill, 2009), the influence of shape and color information on the processing of competitors of the same type (Huettig & Altmann, 2007), and reference tracking of possible antecedents (Dahan, Tanenhaus, & Chambers, 2002). This has been accomplished by varying the target items mentioned within the discourse, or manipulating the referent that is referred to: whether it is a onset cohort, rhyme, or semantic competitor to an item depicted in the visual scene shown with the spoken language (Tanenhaus et al., 1995; Allopenna, Magnuson, & Tanenhaus, 1998; Yee & Sedivy, 2006).

In addition, there have been varying experimental discourse configurations previously studied. Some versions of the paradigm use the discourse aspect to require an action on the part of the participant (e.g., "Click on the bed."; Tanenhaus et al., 1995; Yee & Sedivy, 2006), others passive listening (e.g., "Eventually the man agreed...then he looked at the piano and appreciated that it was beautiful."; Huettig et al., 2006), and in some cases, an eventual response to a comprehension question. These different discourse-processing requirements have been used to address questions regarding integration and access (Dahan, Tanenhaus, & Chambers, 2002; Tanenhaus et al., 1995; Allopenna, Magnuson, & Tanenhaus, 1998; Huettig & Altmann, 2005; Huettig et al., 2006; Yee & Sedivy, 2006). The VWP was flexible enough to test the use of sarcasm in discourse, when considering previous prosody manipulations, and the addition of a question or movement task at the end of a stimulus assured participant attention was on the displays.

The second portion of the VWP that is necessary is the visual display. It is eye-movements to the scenes portrayed in this portion of the paradigm that serve as the dependent variable of analysis. The visual scene itself has taken a variety of different forms in the paradigm since its inception. It has consisted of words typed out (McQueen & Viebahn, 2007), a blank screen while the discourse is played (Altmann, 2004), and most commonly, a display containing approximately four items, of which some are simply distractors and some are related to the spoken discourse (Allopenna, Magnuson, & Tanenhaus, 1998; Huettig et al., 2005; Huettig & Altmann, 2006; Yee & Sedivy, 2006). The placement of the items in the display is usually randomized, and participants are not specifically directed to focus their attention to any particular item. Nevertheless, this can vary according to the instructions given in a particular version of the VWP. Attention is then considered to be a consequence of the linguistic input the participant is receiving that accompanies the display. Although the display contains only a closed set of possible referents, the paradigm is considered to be able to address the questions regarding lexical access and processing through the final portion necessary in the VWP, the linking hypothesis. By altering the items in the display to represent the referents from the multiple meanings activated by the homophones, or items closely related to those meanings, eye-movements to the display offer time course information on the processing of the homophones.

A linking hypothesis (Tanenhaus et al., 2000) is also required for any VWP experiment. Within the linking hypothesis for the study, is the implication that eye-movements to items depicted in the display, when time-locked to the spoken language being presented, provide a measure of language processing that is interpretable for the

research question of interest (Tanenhaus et al., 2000; Allopenna, Magnuson, & Tanenhaus, 1998). Thus, this hypothesis is specific to a set of experimental conditions that are considered valid and reliable by the experimenter. Multiple methods of analyzing the data from this paradigm have recently become available, and are utilized within this work.

CHAPTER 5

GENERAL AIMS

The following sequence of experiments was intended to test the role of sarcasm as affecting ambiguity resolution in discourses, such that it is capable of highlighting a particular alternative explanation when compared to a sincere statement. More broadly, this sequence of experiments was carried out in order to determine whether thoroughly investigating an area of discourse processing sarcasm affects leads to further support for theories of use and resolution already available. The design of the project altered the constraints of the social context for the conversation (and the task) across experiments to determine whether Sarcastic Prosody affected processing to a different extent. The modification of the conversations and the position of the homophones within the sentences was modified such that more or less time was allotted to process the homophone referent after lexical access, and it was spoken in Sarcastic Prosody or not. Specifically, the project sought to determine whether sarcasm that is presented following or simultaneously with an ambiguous homophone referent affects the ongoing discourse interpretation and final resolution of the reference. Additionally, the project aimed to investigate the interaction of sarcasm processing with other characteristics of the discourse contributing to the overall context of the utterance. By examining these factors, we can begin to address whether we can define parameters for an optimal context for sarcasm to be used in, when arriving at a particular discourse interpretation is the goal.

The discourses within the eye-movement studies central to this investigation (Experiments 2 & 3) all contained an ambiguous homophonic referent, which in terms of Bias (Dominance) was either Balanced (two meanings equally likely) or Biased (with Dominant and Subordinate meanings), which themselves were either High or Low Frequency. The homophone references were placed in conversations within VWP experiments, which intended to examine contexts in which sarcasm affected discourse processing when resolving the ambiguity (Tanenhaus et al., 1995). First, a series of norming studies (Chapter 6, Experiment 1) is reported to demonstrate validation of the items and methodology, and quantify covariates. Two VWP experiments are also reported, the first (Chapter 7, Experiment 2) in which a sarcastic utterance followed a homophonic reference in a second sentence, and a second (Chapter 8, Experiment 3) in which the same speaker used Sarcastic Prosody to mention a homophone. The two designs were intended to create different social constraints and contexts with the conversation, leading to different expectations on the part of listeners. Listeners were intended to process the sarcasm in Experiments 2 & 3 while having different expectations of the purpose of the sarcasm in the conversation. These different social situations and the expectations are explained in more detail in Chapters 7 & 8.

CHAPTER 6

EXPERIMENT 1: NORMING STUDIES

The first series of experiments served the purpose of 1) verifying the items created using the homophones would work within eye-movement studies (Experiments 1A-1C), and 2) quantifying the covariates to be included within the analyses of the eye-movement data (Experiments 1D & 1E). Experiment 1A aimed to verify that both meanings of the homophone were activated when it was spoken. The second experiment sought to verify that the visual displays used within the eye-movement studies accurately portrayed the contextual representations for each homophone meaning (1B). The third experiment as aimed to verify that the auditory prosody manipulations worked (1C). Within the study of the covariates, Experiment 1D sought to identify any visual display bias that may exist; while the final experiment (1E) recorded Social Contextual Ratings for consideration. Experiments 1A-E, were therefore carried out before the items were used in the eye-movement studies presented in the following chapters. Below, each norming study is individually motivated and described, after a common materials section.

Materials for Experiments 1A-1E

The materials for the series of norming studies presented here were derived from the discourses created for the eye-tracking experiments, and as such, their transformation for use in each of the norming experiments will be described in detail in the assigned experiment. Thirty homophones were used within the experiment sequence, and the list can be found in Appendix A. This list includes frequency estimates for each taken from

Kucera & Francis (1967). There were 15 ambiguous words (same orthography, e.g., "bank") and 15 homophone pairs, where the orthographic representation did not correspond (e.g., "flower" vs. "flour"). It was concluded that if there were no differences between these groups in the current experiment with the auditory presentation experiment, the groups would be collapsed for further data analysis (and they were). As noted in a previous chapter, using both ambiguous word pairs and homophone pairs allows for a collection of accurate word to meaning corpus frequency data (for the homophones with different orthographic representations) for a portion of the items, in addition to the dominance ratings. Frequency estimates corresponding to the meanings of ambiguous words are noted to be more difficult to obtain since it is not always possible to distinguish between the uses of multiple meanings in a corpus (however, bias estimates pertaining to dominance were calculated for all items). The use of both the ambiguous words and the homophones allowed for a test of the Frequency class effects, while still including enough items for the main experiments, which required a more restricted item type, due to the properties necessary of the stimuli. This included restricting the pairs to items which could have a visual contextual depiction associated with each meaning, and only using homophones for which dominance ratings were available from the population of interest. Items also needed to be constructed so that plausible, neutral situations could be created that fitted both meanings of the homophone, and that semantic associates could be obtained for both meanings to verify their activation. The full discourses for each of the eye-movement experiments can be found in Appendix B.

6.1 Verification Studies

Three verification studies were carried out to ensure the characteristics that made the homophones planned for use in the eye-movement studies appealing, were indeed processed by participants within a spoken language paradigm. This included ensuring that both meanings of the homophone were activated when it was spoken aloud (Experiment 1A), the pictured contexts were appropriate for the meanings (Experiment 1B), and that participants recognized the prosody manipulation (Experiment 1C).

6.1.1 Experiment 1A Introduction

In order to verify that both meanings of the homophones were activated upon hearing it, a cross-modal priming lexical decision task (hereafter LDT) was used, in which participants heard a sentence, and were required to respond to a visually presented word upon completion of the sentence. In this design, the homophone selected was the last word in the sentence, with the sentences constructed from the eye-movement discourses. Automatic semantic priming to written text is well documented in the lexical decision paradigm (Meyer & Schvaneveldt, 1971; Posner & Snyder, 1975; Neely, 1991). In both early and current versions of the semantic priming lexical decision task (Meyer, Schvaneveldt, & Ruddy, 1972; from Joordens & Becker, 1997), the word that the participant makes a response to either shares some form of relation with the word preceding it (meaning-based relation, associative relation, or both), or is unrelated to the target word (Meyer & Schvaneveldt, 1971). In this paradigm, participants are found to respond faster that a target is a word, or that both are words, when it is preceded by or presented with a related word, than when it is not (Meyer & Schvaneveldt, 1971; Neely, 1991; Lucas, 2000; Hutchinson, 2003). These priming effects are typically noted in

terms of changes in response times that differ by milliseconds, when compared to a "neutral" or unrelated baseline word. Because the current experiment set focused on the comprehension of spoken language, the cross-modal aspect ensured that the spoken language was capable of generating a facilitative priming effect. Performing this task as cross-modal priming rather than a traditional lexical decision task also allows it additional comparison to the VWP, which itself has been described as a "visual semantic priming" paradigm (Huettig & Altmann, 2005).

An important aspect of task design within the cross-modal lexical decision task is the effect of relatedness proportion (de Groot, 1983; de Groot, 1984). Relatedness proportion is the proportion of items in the study that share a semantic relation in the priming task, compared to other item pairs, such as the frequency control and unrelated baseline conditions. Previous work established that participants are more likely to engage in strategic processing of primes if long stimulus onset asynchronys are employed and there is a high ratio of related items. Increased RP in item sets appears to be identified by participants, and encourages strategic processing with a sufficiently long stimulus onset asynchrony, a conclusion supported more recently by Lucas (2000) and Hutchinson (2003).

6.1.2 Experiment 1A Method

Participants

For each experiment presented within this series (Experiments 1-3), different participants were selected. That is, no participant was in more than one experiment. A total of 36 participants were recruited for Experiment 1A from the University of South Carolina Department of Psychology's undergraduate participant pool (n = 9 per list

described below), and received course credit when applicable in exchange for their participation. The experiment lasted approximately 10 minutes.

Materials

For the current experiment, only the discourses from Experiment 2 were used, which contain the ambiguous reference to the homophone in the first sentence (since the goal was to ensure both senses were activated). In addition, the sentence was always presented in a sincere prosody. The primes for the current experiment consisted of the homophones, while the targets were created by using the University of South Florida Association Norms (Nelson, McIvor & Schreiber, 1998) for the majority of the homophones present in the norms (and taking the first item that could be pictured and was related to the correct interpretation of the homophone chosen for the study). In the case where it was an ambiguous word, the first word that related to each of the intended senses of the word was chosen. The remaining 8 associates were chosen by using Internet search suggestions of related topics, as they were not present in the norms. The baseline condition of the study consisted of unrelated words, as the use of neutral baselines has become controversial (whether "neutral" baselines such as *blank* or *XXXX* are indeed "neutral" or encourage repetition priming, or other unintended priming effects) (Balota & Lorch, 1986; McNamara & Altarriba, 1988; McKoon & Ratcliff, 1992; from Hutchinson, 2003).

The neutral control words were matched to the targets' reported frequency in the printed frequency of category of information given by the University of South Florida Association Norms. Since this category was derived from the Kucera & Francis (1967) printed norms, the unrelated words were matched in frequency and word length from the

same corpus. This created four experimental categories for each homophone prime (e.g., targets related to first and second meanings, and neutral frequency-matched target to first meaning target, and neutral match to second meaning target). Each participant saw only one word in one condition. See Table 6.1 for a selected item in all possible experimental conditions. The targets were presented visually immediately after the last word of the auditory sentence (the prime) with no variable stimulus onset asynchrony to offset the likelihood of strategic processing, such as that initiated by a higher relatedness proportion. Previous work using ambiguous words has indicated that the multiple meanings are activated quickly (and then one settled on), so the effects of interest here are more likely to be captured at this early time (Seidenberg et al., 1981).

Table 6.1 Sample prime with all associated targets from Experiment 1A.

Condition	Prime	Target
1st Meaning	Bulb	Lamp
2nd Meaning	Bulb	Tulip
Unrelated 1st Meaning	Bulb	Mist
Unrelated 2nd Meaning	Bulb	Isle

Each of the discourses in Appendix B consisted of two sentences that were spoken by two different speakers for Experiment 2, a manipulation that was balanced across the items, such that participants in the current experiment (and every one after with the items) heard the same number of sentences from a male and a female.

Since the homophone of interest was always presented in the Sentence 1 portion of the items, the Sentence 2 portion of the discourse from Experiment 2 presented in

sincere tones served as fillers, with the targets from other experimental items not used in the current list (as there were 4 rotating lists balanced) used as targets for an additional 15 "word"-filler trials, and non-words created for the remaining 15 primes. Thirty additional fillers were also created as non-word trials in order to maintain an equal word:non-word ratio. The non-word targets in the conditions were matched to the experimental targets in terms of average length of characters. Non-words were created by entering in search constraints in the Washington University English Lexicon Project website (http://elexicon.wustl.edu) (Balota et al., 2007). The average length of the target for the experimental items was 5.02 (*sd* = 1.47) characters across conditions, and the character length restriction for the non-words was set for 4-8 characters. In addition, non-words were also selected so as to be pronounceable, and to have a mean reaction time in the corpus ranging from 500-1500 *ms*. From the results returned, 45 non-words were chosen for the current experiment, such that the average length was 5.20, *sd* = 1.34, to approximate the experimental items.

Thus, the 30 experimental items were therefore mixed in with a total of 60 fillers, with each item appearing in one condition per subject. This design allowed the maintenance of a relatedness proportion of 0.167 within the experiment, with experimental items having an associated target for one random sense (15/30), an unrelated word matched in frequency to one random sense (15/30), presented visually on the computer screen for a lexical decision following the homophone offset at the end of a sentence heard aurally, such that an item appeared in all four conditions (both unrelated meanings, both frequency matched controls) across the 4 aforementioned lists, and each participant responded to the same number of related and unrelated experimental prime

and target pairs. Filler items were followed by a non-word (n = 45), or an unrelated word (n = 15) in order to balance the word:non-word ratio, as well as the related:unrelated ratio. The unrelated words were also derived from the Washington University English Lexicon Project website (http://elexicon.wustl.edu) (Balota et al., 2007) and were created by specifying the Kucera & Francis (1967) frequency from 1-175 (to approximate the experimental conditions) and a length count of 4-9 to create the list. The average frequency resulting from this was 83.87, similar to the average for the experimental items. This design follows a similar cross-modal priming design used by Braun & Tagliapietra (2010), but presents the prime word as the last word in a sentence, rather than individually, before it is immediately followed by a visually presented target word to which participants indicate their choice. The experiment was administered using E-prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002).

Procedure

Participants arrived at the lab, and upon indicating informed consent were placed in an isolated room. Participants began the priming experiment with short practice session, before continuing to the experimental items. They heard a sentence, and then made a lexical decision in response to a word presented on the computer screen by pressing two keys on an E-prime button box, one for "yes" and a second for "no" to indicate their answer.

6.1.3 Experiment 1A Hypotheses

It was expected that both intended meanings of the prime would become activated upon hearing the prime in the sentence (and the individual meanings are referred to as 1st meaning, and 2nd meaning in the analyses for all homophones), and that priming effects

will be observable via faster response times for the target words related to the prime, than nonrelated word controls. For unbalanced homophones, this was expected to be the case when the prime was followed by a target corresponding to either the dominant or subordinate meaning of the prime.

6.1.4 Experiment 1A Results

All analyses for Experiments 1, 2, & 3 were carried out using the R statistical software package (v.3.0.1, (R Development Core Team, 2012)). First, the data were analyzed for accuracy. Subject accuracy overall was high, ranging from 87-100%, and no individual subjects were removed due to low accuracy. Overall item accuracy within the experiment was 95%. When the fillers were removed, accuracy for the experimental items and their controls (the Targets of interest) rose to 96%. Only the reaction time data for which correct responses to trials of interest were made were used within the analysis, which excluded another 40 trials across participants. The data were further screened for outliers, with 4 responses being removed for being over 2000 *ms*, and 6 removed for being below 400 *ms*. The average reaction time for the four conditions of most interest (both meanings of the targets, which depending on the item, corresponded to the balanced meanings or the dominant and subordinate meanings, with the length & frequency matched controls for both) was 726 *ms* (*sd* = 245 *ms*). The large variability among response times was examined in a box plot, from which the cut offs were determined. These procedures created a total data loss of 4.7%.

In the current experiment (and Experiments 2 & 3) the lme4 mixed-effects models package of R was used to calculate mixed-effects models for the data (Bates, Maechler, & Bolker, 2011). The mixed-effects models included random intercepts for participants

and items. ICCs were calculated via null models containing only participants (ICC = .31) and items (ICC = .03). These were initially calculated to demonstrate the importance of including the random effects throughout the models in the current work. There were significant priming effects as expected, providing an overall validation of the items for use in the visual world studies. When compared to their baselines, both the 1st and 2nd meanings of the items showed overall priming effects. There was no difference between the priming times for either target related to the meanings of the homophones, and both were significantly faster (~60 *ms*) than length and frequency matched controls. The model tested can be viewed in [4]. All model notation used within the current work is taken from Starkweather (2010).

[4] $Y_{ij} = \beta_{0i} + \beta_1 * \text{PrimeCondition}_{1ij} + b_{i1} * \text{Subject}_{1i} + b_{i2} * \text{Item}_{2j} + \varepsilon_{ij}$

This method was used as an alternative to traditional *F1* & *F2* analyses as recommended by Baayen, Davidson, & Bates, (2008) & Baayen (2008). Presented in Table 2 are the fixed effect model results for the 1st meaning (e.g., "light" bulb) target response times and the difference from the control groups. The response times were grand mean-centered, and therefore the coefficient estimates, including the intercept reference group of the priming for the first meaning, reflect the shift from the grand mean. The model was also tested with the 2nd meaning targets as the reference group (e.g., "flower" bulb), and this model too showed that the control conditions were significantly slower, and that the two meanings were indistinguishable from one another in terms of priming. Thus, the results of Experiment 1A verified that both meanings were reliably activated across experimental items.

Table 6.2 Results for Experiment 1A.

Coefficient estimates for fixed effects for mixed-effects model where reference group included response time to the first related meaning. P – values were calculated using Markov chain Monte Carlo estimation. Significant effects in bold.

Condition	*Est.*	*Std. Error*	*t*	*p* <
Intercept (1st Meaning)	-81.527	27.33	-2.983	**0.001**
2nd Meaning	5.977	17.25	0.346	= .73
Control to 1st Meaning	58.877	17.46	3.371	**0.001**
Control to 2nd Meaning	70.736	17.42	4.060	**0.001**

6.1.5 Experiment 1B Introduction

The second norming study was intended to verify that participants were able to recognize the context depicted by the visual displays, which consisted of photos of the contextual depictions of the homophones. In Experiment 1B, participants ordered photos of the contexts (out of a possible 3) by typicality. For an example of the screen presented to participants for "bulb" (flower), see Figure 6.1. Participants were also given a word to disambiguate which meaning (e.g., given bulb, they were given "flower" in one case) was meant by the contexts for ordering. These measures of typicality were then used to select the final image for inclusion in the eye-movement studies. A more detailed methodology is presented below.

Figure 6.1 Sample visual display for "flower" screen for bulb in Experiment 1B.

6.1.6 Experiment 1B Method

Participants

Thirty different participants were recruited from the University of South Carolina Department of Psychology's undergraduate participant pool and given course credit in exchange for participation. The experiment lasted approximately 10 minutes.

Materials

Each participant ordered the typicality of a series of 60 image sets, in which there were 3 images per set. The 60 image sets corresponded to the contexts for the homophone pairs to be used in Experiments 2 & 3, and were gathered from various online outlets and photographs by S.A.P. Each photo was cropped to the same size for inclusion (a 6" x 10" variant). Each photo contained a contextual depiction of a homophone, in which visual cues give representations of the intended meaning. The

most typical context representation, as chosen by the population of interest, was then used for study in the remaining eye-movement experiments, with the pictures chosen randomly placed in the visual display for the eye-movement studies. The experiment was administered using E-prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002).

Procedure

Participants arrived in the lab and upon indicating informed consent, were seated in isolation. They were told they would be presented with a context, and asked to order the images also presented by typicality, with their first selection corresponding to the most typical image corresponding to the label given. With each set of pictures, participants were given the word that the context was intended to depict, as well as a word that disambiguated the intended meaning (e.g., "bulb", given flower). A display signaled the end of the experiment.

6.1.7 Experiment 1B Results

The photographs selected as being most typical by the agreeing majority of participants in the norming study were used in Experiment 1D, and later in Experiments 2 & 3. The goal of this experiment was to verify that the contexts to be used for the study are both recognizable and typical of the homophone. The most frequent choice among the 30 participants (trial presentations were randomized in the placement of each possible picture within the display across participants, as well as the order of presentation) was selected for use. The range of selection proportion from the 3 choices that landed the target as the chosen option ranged from .367 to .567. In other words, pictures chosen to represent the context had to have a majority proportion of the vote. Two ties were broken

by the experimenter selecting a random picture from the set (in both cases, all 3 pictures received 10 votes, or were considered equally representative).

6.1.8 Experiment 1C Introduction

A third norming study was completed before the eye-movement studies were conducted in order to verify that an audio manipulation of the sincerely recorded sentences did in fact make them sound sarcastic to the population of interest. The manipulation had been used in previous work (Peters, Wilson, & Almor, 2012) and was derived from work identifying auditory cues interpreted by participants as "sounding" more sarcastic (Rockwell, 2000; Cheang & Pell, 2008). In order to verify the current items, the manipulation was carried out and then the items were normed drawing from the same undergraduate population as previous studies utilizing the manipulation.

6.1.9 Experiment 1C Method

Participants

Two separate norming experiments were carried out (one for Experiment 2 discourses, one for Experiment 3) with participants from the University of South Carolina Department of Psychology's undergraduate participant pool (n = 9 for Experiment 2 items, n = 10 for Experiment 3 items). Participants received course credit when applicable in exchange for their participation. The experiments lasted approximately 10 minutes each, with each participant only completing one.

Materials

Individual sentences for conversations in Experiments 2 and 3 were recorded by native English female (S.A.P.) and male speakers (T.W.B.). During the two experiments, the sentences were very similar, and the male and female speaking parts (either Sentence

1 or Sentence 2) were fixed across experiments (e.g., if the male speaker had Sentence 1 in item 1 in Experiment 2, they had the same item in Experiment 2). The 2-sentence discourses that were created to be used in eye-movement experiments (2 and 3) were recorded in their entirety using a sincere tone. Sentence 2 sincere items were edited by introducing Sarcastic Prosody cues to create the sarcasm condition (e.g., the cues of tempo, intensity, and pitch previously described in the introduction). This was done separately for each item using Praat software (Boersma, 2001) in the following manner.

A previous study provided the baseline for natural Sarcastic Prosody cues using Praat (Peters, Wilson, & Almor, 2012). The previous experiment used the identification of the main areas of difference noted when sincere statements were compared to the same statement said sarcastically had lower pitch, slower tempo, and slightly higher intensity than in the sincere versions. The differences noted in the original comparison were used as a rough model (further refined by trial and error, particularly relating to the amount of the phrase that was manipulated) for the adjustments made to the manipulated copies of the sincere versions of the sentences.

To begin, in each sentence, an area of manipulation was identified. For the reported speech phrase, pitch was adjusted to about 0.9% of the original, and duration increased by 30% (in some cases, the duration of the entire sentence was edited, again, this was done on a trial and error basis). Intensity was increased and contoured such that the stressed syllable of the first stressed word of the phrase (in multiple word phrase items) was given a 2.5 db multiplier, which was sloped down to a 2 db multiplier at the end of the sentence, or phrase as needed. For a key word, duration in some cases was again increased again by 30% on the stressed syllable. As a result, in the current

experimental design, functionally sarcasm is operationally defined as speech having the manipulations of "lower, slower, and louder". In Appendix C, spectrograms of a "Sincere" and "Sarcastic" item can be viewed. The ability to differentiate between sincere and sarcastic statements using prosody information resulting form this manipulation was previously demonstrated in the same population of participants to be successful, with an estimated Cohen's $d = 1.02$ (Peters, Wilson, & Almor, 2012).

In the current experiment versions, participants heard a total of 150 sentences. The sentences were randomized for presentation for each participant. To sum the 150 sentences, they heard Sentences 1 and 2 of the filler items (60 sentences total), as well as the sincere and sarcastic versions of Sentence 2 of the experimental items (60 sentences total, 30 of each), and Sentence 1 of the experimental items (30 sentences) for Experiment 2 or 3.

Procedure

After arriving at the lab and indicating informed consent, participants were taken to an isolated room. They were instructed to rate the sentences that they heard on a scale of 1-5 for sincerity, with 1 being the most sincere, and 5 being insincere (the scale and the meaning was presented during each trial). They were also told that they would hear some sentences twice, and to respond to the version of the sentences that they heard.

6.1.10 Experiment 1C Results

One subject was removed from the rating of the Experiment 2 items for reversing the scale, leaving $n = 8$ within the analysis. In addition, only half of the rated fillers were included in the analysis in order to balance the conditions ($n = 30$ for each). The filler ratings kept in were not significantly different from those left out for either experiment

(t's = 1, p > .05). The results of the rating studies are reported below by experiment and condition in Table 6.3. The difference in the rating conditions for each experiment was verified for the results using a within-subjects repeated-measures analysis of variance (*F1*) and a between-subjects items analysis of variance (*F2*). In the ratings for Experiment 2, $F1(3, 21) = 157.60$ and $F2(3, 956) = 146.91$, p's < 0.001. After a Bonferroni correction for multiple comparisons, differences were found between the conditions, such that the Sarcastic items were rated greater than all other conditions, the Sincere items were not rated as different then the Sentence 1 items, and the Fillers were rated slightly lower than the sincere items and Sentence 1 items. The same analysis of Experiment 3 found $F1(3, 28) = 117.60$, $F2(3, 1083) = 112.51$, p's < 0.001. The same adjustment found once again Sarcastic items rated higher than the remaining conditions. The Fillers were rated similarly to Sentence 1, while the Sincere items were rated slightly higher than both the Fillers and Sentence 1 items. Thus, in both experiments the Sincere items were rated lower than the Sarcastic items, as intended by the manipulation.

Table 6.3. Rating means and standard errors by Experiment and Condition.

Experiment Items From	Discourse Condition	Mean (*Standard Error*)
2	Sentence 1	2.542 (*0.073*)
2	Sentence 2 Sincere	2.675 (*0.082*)
2	Sentence 2 Sarcastic	4.213 (*0.072*)
2	Fillers	2.158 (*0.072*)
3	Sentence 1	2.083 (*0.068*)
3	Sentence 2 Sincere	2.780 (*0.080*)
3	Sentence 2 Sarcastic	3.743 (*0.077*)
3	Fillers	2.127 (*0.068*)

For the Experiment 2 items, based on the ratings, 2 of the sarcastic items needed to be re-manipulated to sound more "sarcastic" overall (items 24 & 30), as well as 1 of the Experiment 3 items. A separate norming experiment containing the versions of the 3 problematic items was conducted with an additional 2 participants. The norming study itself was an abbreviated version of the previous studies, simply to verify that the problematic items were fixed with the re-manipulation. Two of the items, (24 and 30 from Experiment 2) were outright successfully remanipulalated, with higher average sarcastic than sincere ratings. One item in Experiment 3, (item 4) was still problematic due to the small number chosen for the rating, and when the analyses for Experiment 3 are reported, they are reported without the item, (for item 4, one participant tied the ratings for sincere/sarcastic versions, one had them reversed). The experiment series then proceeded with the manipulated items.

6.2 Covariate Analysis

The final two experiments addressed possible covariates for use within the eye-movement studies. The first experiment aimed to measure characteristics of the visual display. The second aimed to measure pre-existing bias for social context for the given homophones.

6.2.1 Experiment 1D Introduction

A forth experiment was carried out in order to ensure that the photographs chosen for each item lacked inherent bias, caused by factors such as one being more visually dynamic than the other, etc. In order to ensure this, or account for effects if they were there, an eye-movement study in which participants listened to short sentences (which were unrelated to the passages) while viewing the visual displays was carried out. These

sentences were filler items from Experiment 1A, and special attention was paid such that they were unrelated to the pictures.

6.2.2 Experiment 1D Method

Participants

Thirty different participants were recruited from the University of South Carolina Department of Psychology's undergraduate participant pool, and received course credit when applicable in exchange for their participation. The experiment lasted approximately 5 minutes.

Materials

The two visual displays rated as most typical by the majority of participants in Experiment 1B were combined to form the figure for each item in Experiment 1D. That is, the most typical flower shop and hardware store were combined to form a display as shown in Figure 6.2, for example, for the word "bulb".

Figure 6.2. Sample experimental visual display for Experiments 1D, 2 & 3.

After the visual display pairs were completed, the verbal filler items that were used as additional fillers in Experiment 1A were randomly matched with the pictures, with the only constraint being that no contexts overlap. The resulting 30 items were used to create a short VWP experiment to assess any differences in the appeal of the pictures, to be later factored into the analyses of Experiments 2 & 3. Participants were asked to listen to the sentences as well, and after a picture set, to determine if a word presented on the screen was heard or not. In half of the trials the word presented was present, in half it was not. Also, there were two versions of the display, such that some participants saw one picture on the left and the other on the right, and another saw the reverse. These

were later combined for analysis, and the pictures were placed randomly in Experiments 2 and 3.

Apparatus

Participants' eye-movements were recorded using a desktop-configured with chinrest SR Research Eyelink 1000, which sampled eye position at 1000 Hz (SR Research Ltd., 2013). Accuracy with this model is reported to be within 0.25° – 0.5° on average. Eye-movements of the right eye were recorded, and the visual stimuli were presented on a 19" CRT monitor positioned ~ 47 cm directly in front of participants. The experiment was controlled by SR Research computer running Experiment Builder software (SR Research Ltd., 2013).

Procedure

Participants arrived at the lab, and after indicating informed consent, heard instructions from the experimenter, and read them on the screen. Before beginning a short practice session, the experimenter calibrated the tracker for the participant using a 9-point calibration procedure. Calibration was repeated during the experiment as necessary. The instructions for the experiment were given as follows.

Participants were told that their task was to listen to the sentences and be able to respond to whether a word was present or absent during the sentence. Participants were asked to indicate their responses to a word being present by pressing "1" for "yes" or "2" for "no" after it appeared on the screen. They were also told that they needed to keep their eyes on the pictures at all times, but not given any instruction as to where to look, other than to naturally observe the pictures. They were told that there would be a short recognition task after the current experiment where they would be asked to identify some

of the pictures present in the experiment from distractors. After indicating their answer to the word probe, the next visual display appeared, and the next corresponding auditory item began.

Previous work has demonstrated that tasks are necessary within VWP studies in order to assume that the linking hypothesis being used in the experiment is in fact accurate, and that some tasks are better within this paradigm than others (Salverda, Brown & Tanenhaus, 2011). While the suggested norming study here does not implement what some would consider to be the most advantageous task design, the goal was to evaluate baseline eye-movements based on interest levels, and for that goal, the design was acceptable. Experiments 2 & 3 involve the use of goal directed behaviors that are recommended for use in this paradigm (Salverda, Brown & Tanenhaus, 2011).

6.2.3 Experiment 1D Results

Data Preparation

Eye-movement data for each trial was collected by Experiment Builder (SR Research Ltd., 2013). Pictures were assigned to the correct condition by identifying the version of the Experiment (A or B) that the participant had been in, and treating their data accordingly. Later, fixation and dwell reports were generated for the experiment that allowed the calculation of where participants were fixating throughout each trial, to be assigned according to the pictures on the display. Areas of interest within the visual display were designated by picture in the analyses. The results generated included fixations from the full length of time the picture was available on screen, since identifying any effects generated by the picture was the focus.

After the data were prepared, average overall dwell times for each picture during the course of the unrelated sentence were calculated, as well as the first fixation duration for each. Overall dwell times are considered to reflect the continuous processing of the item throughout the trial, while first fixation measures are typically viewed as a measure of initial processing of a stimulus.

If there were no inherent differences in the visual displays of the items, then there should be no overt preference for items in the display. Overall, the average difference across item pair in terms of total dwell time was ~24 *ms* (calculated by looking at the difference between the average overall dwell time of Picture 1 vs. Picture 2 for each item). Overall average dwell times ranged from 1122-1901 *ms*, but there was high variability amongst items, as the average of the standard deviations across items of the dwell times was *sd* = 504 *ms*. In terms of average first fixation duration between item pairs, overall there was little difference in length, ~16 *ms* (calculated by looking at the difference between the overall average first fixation time of Picture 1 vs. Picture 2 for each item). First fixation durations for individual pictures ranged from 202-438 *ms*, and the average standard deviation was lower, *sd* = 142 *ms*.

The results of this experiment indicate a high degree of variability by item and participant. In Experiments 2 & 3, participants were given additional time to preview the display before the audio items begin to play, rather than having the two events occurring simultaneously. This was intended to address any first fixation differences. A proportion of overall dwell time on each picture was created to be considered as a covariate in Experiments 2 & 3, calculated by taking the average dwell time for each picture, and dividing it by the total average dwell time for the two pictures, for each item.

In addition, in order to ensure that the effects of interest within the eye-movement studies could be linked directly to the offset of the *mention* of the homophone of interest, rather than just its contextual depiction, analyses were performed to confirm that the effects of Dominance and Frequency were not present when just viewing the picture. A two-factor analysis of variance was calculated with the overall proportion of dwell time for each picture (calculated by taking the average dwell time for each picture, and dividing it by the total dwell time on the two pictures) as the dependent variable and factors Dominance (Subordinate vs. Dominant vs. Balanced) and Frequency (Low vs. High) per picture. Neither factor had a main effect, and the interaction was insignificant, F's < 1, $p > .05$. A similar two-factor analysis of variance with mean number of fixations for each picture determined by Dominance and Frequency also yielded no significant main effects or interaction of either factor (F's < 1, $p > .05$).

In summary, this experiment demonstrates that across the visual displays, there are differences in preference, but they do not appear to be systematically related to characteristics of the homophone depictions, such as written corpus Frequency or Dominance measures. However, one final analysis correlated the overall proportion of dwell time for each picture with the average amount of priming for the item calculated in Experiment 1A, and found a significant medium strength correlation between the two measures ($r(58) = 0.298$, $p = .021$), such that the items receiving a greater proportion of fixations during the current experiment, also received more priming within Experiment 1A. This suggests that the dwell ratio, rather than being dependent on a particular characteristic of the homophone being depicted (such as its Frequency or Dominance), may be related to the strength of the overall processing of the meaning depicted, which

may sum these cues, as well as other characteristics of the homophone, to result in these overall processing measures.

6.2.4 Experiment 1E Introduction

A second study of potential covariates was completed in which a Social Contextual Rating was calculated for each member of the homophone pair. The social component of sarcasm has been noted in the literature as well as a discussion of individual differences (Happé, 1993; Sullivan, Winner, & Hopfield, 1995; Chevallier et al., 2011; Colston & Lee, 2004). This was an attempt to control the social and context component, owing to the fact that some of the homophones (e.g., "hair") are more likely to lend themselves to being part of a sarcastic comment. The idea for this work was also previously introduced in the discussion of the theoretical approaches. By asking the participant population directly whether they think a particular meaning of a homophone is more tractable for an expected sarcastic comment, the results from this study were able to serve as a secondary measure of sarcasm, or place a measure of sarcasm on a continuum rather than the "all or nothing" measure the prosody manipulations contributed. If in the data, a sarcasm prosody manipulation (all or none) factor is not sensitive enough to capture the effect, the social contextual knowledge provides an alternative measure of sarcasm. If there is no interaction between the measures, it could be suggested that the social contextual knowledge overrides or encompasses the prosody information when it is available.

6.2.5 Experiment 1E Method

Participants

Thirty participants were recruited from the University of South Carolina Department of Psychology's undergraduate participant pool, and received course credit when applicable in exchange for their participation. The experiment was conducted by an online hosting website (sc.sona-systems.com; (Sona Systems, Ltd., 2013)) and was self-paced. Participants finished in approximately 5.5 minutes (*sd* = 392 *ms*).

Materials

Participants were presented with a word (each homophone) and a word that disambiguated what meaning was intended (e.g., oar and "paddle"). They were then asked to rate how likely they would be to make a social comment for each of the items on a scale of 1-5, with 1 being "not likely at all" and 5 being "very likely". Hypothetical examples were also given, as reported in Table 6.4.

Table 6.4 Sample "how-to" instructions for online social norming study.

Participants were told that they may encounter the same word twice, and to respond to the meaning of the word that was presented with the word.
1. How likely are you to make a critical comment that involves the meaning of the object? STRAW (Drinking) For example, you may think a drinking straw is likely to be used to drink soda and as a such you may consider that it could reasonably evoke a critical social comment, for example about someone blowing soda using a straw. If this is something you consider then you might want to choose a rating of 4 or 5.
2. How likely are you to make a critical comment that involves the meaning of the object? STRAW (Hay) Following the previous example, you might think that straw is likely to be eaten by animals and therefore not very likely to be involved in a critical social comment unless you really stretch your imagination and think about a person chewing straw on a farm. In this case you might want to consider a rating of 1 or 2.

Procedure

Participants were able to sign up and complete the study through the sc.sona-systems.com online hosting website (Sona Systems, Ltd., 2013). Their participation in the experiment was considered to be indicative of their informed consent, as they were able to withdraw at any part during the experiment. After reading the instructions and viewing the examples (presented in Table 6.4), participants were free to complete the task and received credit automatically.

6.2.6 Experiment 1E Results

Rating scores were recorded for each item, and the averages were collected and added as a possible covariate for the eye-movement studies. The ratings across items ranged from M = 1.83 - 3.87, with *sd*s ranging from 0.17 - 0.29. Once again, one way to consider the ratings (when considering the instructions for the experiment) is as a

continuous measure of sarcasm likelihood. A full list of the ratings per item is available with other item information in Appendix A.

Additionally, similar to the analysis in Experiment 1D, analyses in the current experiment aimed to determine whether the Social Contextual Ratings were capturing information that took into account word knowledge (such as Dominance and Frequency homophone information) with sarcasm likelihood, and a two-factor analysis of variance (factors Dominance and corpus Frequency) was calculated to determine whether the factors contributed the social contextual ratings. Within this analysis, there was a main effect of both Dominance $F(2, 54) = 10.568$, $p < 0.001$, and Frequency $F(1,54) = 9.572$, $p = 0.003$. The interaction did not reach significance ($p = 0.081$). Post-hoc testing revealed significant differences in the Dominance factor to be between the Dominant ($M = 2.76$, $se = 0.14$) and Subordinate conditions ($M = 2.17$, $se = 0.08$), and the Balanced ($M = 2.90$, $se = 0.10$) and Subordinate conditions only. Within the Frequency factor, High frequency words received higher social contextual ratings ($M = 2.98$, $se = 0.12$) than Low frequency words ($M = 2.57$, $se = 0.09$).

Thus, it appears that participants engage in lexical processing of the homophone meanings to arrive at the Social Contextual Ratings. In Experiments 2 and 3, eye-movements are likely to be influenced early on by these factors, and this must be considered within the linking hypotheses. Additionally, these factors appear to influence lexical processing not only when a word is presented solely within text, but also when the likelihood of the word being said using sarcasm is evaluated, as was required by the task. This suggests that the information provided by the corpus Frequency and Dominance

ratings will impact the interpretation of the homophones when they are heard, as long as their meaning is evaluated.

Additionally, as a further result of this experiment, an individual differences task was included in Experiments 2 & 3, in order to obtain a measure of each participant's own awareness of social context awareness and appropriate interpretation. This was a Faux Pas task that is described in more detail in Chapter 7. The goal of this experiment was to determine if individuals differ in their show of social awareness of the context remarks are placed in.

6.3 Experiment 1 Conclusions

The results of the item verification and covariate studies allowed the eye-movement studies to proceed as planned. Both meanings of the homophone were initially activated when heard in a spoken sentence (Experiment 1A), and participants indicated that the visually depicted contexts fit the homophones (Experiment 1B). With alterations to individual items, the auditory items were also verified (Experiment 1C). The results of Experiments 1D & 1E were also recorded to be used when necessary as covariates in the analyses of Experiments 2 & 3.

CHAPTER 7

EXPERIMENT 2

7.1 INTRODUCTION

After the verification of the items and measurement of potential covariates in Experiment 1, the remaining two experiments sought to explore the main research questions introduced in Chapters 1 & 5. These questions focus on what happens when we view sarcasm as a piece of information that can contribute to the successful resolution of the ambiguity created by a homophone in a discourse. The effect of sarcasm is considered in relation to the ongoing competing meaning interpretations, and it is hypothesized that sarcasm may serve to highlight a particular alternative. The first experiment presented in the current chapter addresses the question of whether the use of Sarcastic Prosody (versus sincere) in an utterance by a speaker is able to affect the processing and eventual interpretation of an ambiguous referent previously introduced by another speaker. The experiment also considers how aspects of the ambiguous referent, in this case, characteristics of the homophone, interact with the processing of the Sarcastic Prosody during discourse resolution. The second experiment (presented in Chapter 8) further modified the conversation's social contextual setting by combining the homophone mention with the speaker utilizing Sarcastic Prosody, thus affecting the listener's processing expectations as they heard the conversation.

Within the current experiment, no explicit information regarding the proper contextual setting was offered to a participant; instead they were required to build an

affordance for an interpretation of the conversation based on the combination of context and prosody information presented. This information included the lexical information associated with the homophone itself, such as the Social Contextual Rating value (as defined in Experiment 1), Frequency, and Dominance information that was assumed to become available to a listener when the homophone was mentioned (as demonstrated in aspects of Experiment 1); this, in addition to the construction of the phrase that followed and contained the Sarcastic Prosody. While the first three variables have been described within Experiment 1, the last variable was unique to the presentation of the items in the current experiment. The design of the conversation included a first speaker using a homophone to generate ambiguity, and a second speaker following the ambiguity with a Sarcastic or Sincere utterance. By having a second speaker follow the mention of the homophone with a Sarcastic utterance (or not), the expectation is that a listener believes that the second speaker is privileged to the first speaker's intended interpretation, and should therefore pay careful attention to any informational cues offered by the second speaker. This assumption maintains that listeners have knowledge of the "social" aspects of an exchange between two individuals, such as the assumption that they are sharing some conversational common ground, and this expectation is particularly important to consider when examining any effects of Sarcasm by Social Context in the reported findings.

The overarching social context for the reported eye-movement studies that was presented to participants involved that they were hearing short conversations amongst individuals, and they were tasked with determining which visually depicted context (out of a choice of two) best matched the conversation that they heard. By choosing the

picture they thought best matched the discourse (in the experimental items, contextual depictions of homophone meanings were present), participants were forced to indicate how they believed the homophone mentioned, and the discourse, should be interpreted. The resolution chosen was important in order to determine whether the use of sarcasm in an utterance altered final choice (when compared to Sincere conditions), but equally important was the process by which they arrived at the choice, and determining whether the use of sarcasm in an utterance impacts this process. Filler items ensured that the participants were engaged in the task and choosing appropriately by having "correct" contextual choices.

Below, a common Method section is provided, which is shared by Experiments 2 & 3. Beyond the common Method section is a Materials section specific to Experiment 2, including information regarding the items and design of Experiment 2. Within the discussion of Experiment 3 is a similar Materials section describing the items specific to that experiment.

7.2 Common Method For Eye-Movement Experiments 2 & 3

Participants

Sixty participants were recruited from the University of South Carolina Department of Psychology's undergraduate participant pool, 30 for each eye-movement experiment. All received course credit when appropriate in exchange for their participation. Participants were recruited for only one of the experiments in the series, and each lasted approximately 30 minutes. Participants who took part in any portion of Experiment 1 were excluded from participation. In Experiment 2, there were 24 females and 6 males, and in Experiment 3, 23 females, 7 males.

Apparatus

The same apparatus was used as previously described in Chapter 6, section 6.2.2.

General Materials & Procedure

Each of the Experiments (2 & 3) had two parts. After indicating informed consent, participants began the eye-movement study. First, the experimenter calibrated the eye tracker using a 9-point calibration procedure. Calibration was repeated during the experiment as necessary, heuristically when drift correct was in excess of 1° visual degree. The experimenter then gave the instructions for the eye-movement experiment as follows verbally, and the participant could also read them on the computer screen.

Participants were told that their task was to listen to the conversations speakers were having, and respond to the question at the end of each short, 2-sentence discourse. The question always asked which picture they thought best fit the sentences they just heard. After 4 practice items, participants heard a total of 30 experimental discourses in each experiment, mixed with 30 filler discourses. Each discourse was accompanied by a visual display containing two pictures depicting possible contexts for the conversation (see Figure 6.2). For the filler items, the questions and the contexts did not involve ambiguity. At the question screen, participants were asked to indicate their responses by pressing the "1" or "2" key corresponding to the picture on the left or right in the visual display, as soon as they decided on the answer. Reaction time for the choice was collected, with reaction time calculated from the onset of the question screen, to the detection of key press indicating their choice. After indicating their answer, participants fixated at the center of the screen, and the experimenter advanced to the next visual

display, and the next corresponding auditory item began. The eye-movement experiment took approximately 20 minutes for participants to complete.

A second experiment in each series involved a measure to determine the social awareness of the participants, as a result of Experiment 1E. The task was a Faux Pas task in which participants read through ten short passages (4-5 sentences each) and in the experimental items, were required to infer information, such as the social appropriateness of comments made by characters within the passages. In 5 control passages, there was no "social" knowledge required to answer the comprehension questions, while the other 5 passages had comprehension questions that required readers to make an inference when interpreting socially complex situations (e.g., the intentionality of an insult when a person was within hearing distance). Comprehension questions verified whether they had interpreted the passages in the correct manner, or made a social "Faux Pas". The task was self-paced, and took participants approximately 5 minutes to complete. It was adapted from Baron-Cohen & colleagues (1999) and edited for use with American English-speaking adults, as well as adapted for use with a computer. Accuracy of responses was calculated for each participant for the task, and accuracy of their responses in the Faux Pas condition was used as a covariate in a portion of the eye-movement data analyses.

Data Preparation Procedure for Experiments 2 and 3

Eye-movement data for each trial was collected by Experiment Builder (SR Research Ltd., 2013). Later, fixation reports were generated for each experiment independently that allowed the calculation of where participants were fixating throughout each trial to be assigned. Additionally, information from the covariates recorded from

Experiment 1 was merged with the fixation data depending on which picture was currently receiving a fixation (Picture Proportion scores from Experiment 1D, and Social Contextual Ratings for the item, from Experiment 1E). This merged data also included Frequency and Dominance classes for the homophone contexts ("High" or "Low" for Frequency, or "Balanced", "Dominant" or "Subordinate" for Dominance). For Frequency estimates, when the homophone meanings had different orthographies (e.g., "fur" vs. "fir"), direct frequency estimates for each could be used to generate the classification (Kucera & Francis, 1967). When they had the same orthography, the Frequency estimate from Kucera & Francis (1967) was used to generate the classification for both meanings, and the Dominance classification, Balanced vs. Biased (e.g., "Dominant" or "Subordinate") was informative in disambiguating them. This is the reason both variables were used in categorizing each contextual representation of the meaning of the homophone. For the conversations, individual sentence timing information had previously been calculated and recorded using Audacity Software (2010) such that the duration of each sentence, as well as the onset and offset points of the homophone references, were matched to the items and eye-movements within the dataset. Therefore, for each trial throughout the discourse, fixation information was generated according to the specifics of where a participant was looking, and what they were hearing. This allowed for the data models to be time-locked to events of linguistic importance in the eye-movement record.

Analysis Techniques for Experiments 2 & 3

For the data reported, mixed-effects models with random intercept effects for subjects and items were used as an alternative analysis to the traditional *F1* and *F2*

analyses often used in psycholinguistics. The general equation is referenced in [4] in section 6.1.4. The models allow the variability by subject and item to be addressed across the analysis and also allow other relevant covariates to be analyzed within the model when needed (Baayen, Davidson, & Bates, 2008). Thus, they provide flexibility within the analyses, more so than *F1* and *F2* analyses. Within these models, when theoretically supported, interaction effects between covariates and predictors were tested. These interactions are reported within the fixed-effects model data reported for each individual model. When interactions were not included, covariates and predictive factors were entered in the model according to the order presented in the model result tables. All analyses reported in Experiments 2 & 3 therefore had fixed-effects that were significant when subjects and items were included as having a random effect for the intercept, with reported *p*-values estimated using Markov Chain Monte Carlo estimation. Analyses were carried out using the R statistical software package (v.3.0.1, (R Development Core Team, 2012)), and the lme4 software package, which runs the mixed-effects models, (Bates, Maechler, & Bolker, 2011).

7.3 Methods For Experiment 2

Materials

Within Experiment 2, the use of 30 of the 2-sentence long conversations previously verified in Experiment 1 was implemented in order to test whether sarcasm affected the final interpretation of an ambiguous referent, and determine if and when this cue is most useful when visual targets depicting both possible meanings of the homophone are present. As previously noted, in terms of the factor Bias, both Biased (referred to as having Dominant and Subordinate meanings), and Balanced homophones

were used in order to create ambiguity within the conversation as it pertained to a visually depicted scene. Fixations following the onset of the homophone in Sentence 1 were measured within the data in order to verify the lexical information from the homophone was indeed accessed upon hearing it, as a baseline. The visual scenes were created with the results of the picture norming in Experiment 1, such that two pictures were seen on each display, one displaying each possible contextual interpretation. A sample display can be viewed in Figure 6.2. In the example, the ambiguous referent is bulb, with "flower" bulbs possible in the flower shop on the left, and "light" bulbs possible in the context of the hardware store on the right.

In each trial, depictions of contexts for the meanings of the homophones were placed randomly for participants. A sample version of an accompanying verbal stimulus and corresponding question can be viewed in Table 7.1, with all items listed in Appendix B. After each conversation, participants were asked to choose the contextual representation they felt best fit the conversation.

Table 7.1. Experiment 2 Sample discourse.

Sentence	Speaker	
1	A	I feel like I have to buy bulbs every year.
2	B (Sincere/Sarcastic)	Maybe they require more care than you realized.
?	(Text appearing after scene)	Which picture best fit the topic of conversation?

**B presented in Sarcastic Prosody in sarcastic conditions.*

All conversations had two speakers, (one female, S.A.P., one male, T.W.B.) and "speaker" was balanced across conditions such that each speaker had an equal amount of

Sentence 1 and Sentences 2 parts both within the experimental items and the fillers. The design of the experiment included Sarcasm present within an utterance as a factor, where half of the time Sentence 2 contained Sarcastic Prosody. Bias and Frequency were also used within the analysis as factors, such that over the selection of homophones there were a variety of both information types. Due to an error in programming, the proportion of Sarcastic to Sincere items overall within Experiment 2 was .8 (participants heard, on average, 20% fewer Sarcastic items). This error was fixed for Experiment 3, and is accounted for in the analyses of Experiment 2 (within the choice proportion calculations). The design of this experiment is such that it allows an investigation as to whether sarcasm changes the interpretation of the ambiguous homophone by highlighting an alternative (dependent on its associated Frequency, Dominance and Social Context characteristics), and/or aids in resolving the reference one way or another for a listener, when compared to a Sincere statement.

7.4 Linking Hypothesis & Theoretical Predictions

Linking hypotheses assume that eye-movements to the scenes depicted within an experiment provide an accurate representation of the online processing of the spoken items being heard (Tanenhaus et al., 2000; Allopenna, Magnuson, & Tanenhaus, 1998). The predictions from the different theoretical perspectives that follow from the Linking Hypothesis for the current experiment (outlined below), are offered in terms of how the eye-movement and behavioral data should be interpreted. A linking hypothesis is presented separately for each experiment.

The Linking Hypothesis for the current experiment focuses on examining eye-movements at the offset of two key events occurring within each of the spoken

conversation stimuli. These two events include the introduction of the homophone verbally, and the analysis of eye-movements during the Sarcastic Prosody. Eye-movements to the display and their characteristics during both of these events are most likely to reflect processing of the homophone and sentence meaning. Specifically, the linking hypothesis argues sarcasm is hypothesized to have the greatest effect on the ongoing processing of the homophone *as a listener encounters the prosody and is still deciding which meaning is intended*. In this context, sarcasm has the best chance of being seen by listeners as a piece of viable information in resolving the reference. This can be verified by examining eye-movements while participants are hearing the utterance both early and late in Sentence 2, in a Sarcastic versus Sincere prosody. First fixations and average fixation durations were calculated as the dependent variables for both experiments in the windows identified as a measure of immediate attention, as well as shifts in attention through the time windows. The effect of Sarcasm can also be considered by examining the overall sentence interpretation, in which choice proportions per item (one contextual representation versus the second), and reaction times to make a decision can be compared by condition.

The first event, the introduction of the homophone, serves as a verbal description of the visually depicted contexts and occurs toward the end of the first sentence within the experiment. Shortly post-onset of the homophone, it is expected that the lexical representation(s) will be accessed, leading to changes in attention reflected in the eye-movement record as differences in fixation patterns and durations toward the contextual representations. These differences are generated based on which characteristics of the homophone meaning are being accessed, and serve as a baseline in the current

experiment series, for an activation of the homophone meanings with no prosody cue. Previous work has found that multiple representations of the homophone meanings are accessed when it is heard in spoken language (Blunter & Sommer, 1988, from Cutler, Dahan, vanDonselaar, 1997; Grainger, Van Kang, Segui, 2001) (and this was confirmed in Experiment 1), so the immediate effects of the activation of both meanings may manifest in fixations to depicted contexts, but with fixation lengths that differ depending on characteristics of the homophone such as Frequency and Dominance. The relationship between mention and the launch of an eye-movement to the depicted target is well reported within the VWP literature (Yee & Sedivy, 2006; Yee, Overton, & Thompson-Schill, 2009).

The second event of importance, which pertains more directly to the linking hypothesis, is the introduction in the auditory item in Sentence 2 of sarcasm, and the processing that follows. The sarcastic manipulation affects processing of Sentence 2 as a whole, when considering it alters the length of the statement. In the current experiment, sarcasm is introduced in the first portion of Sentence 2, and by examining the first and second 2000 *ms* separately, it can be determined if characteristics of the homophone/sarcasm interaction change as the utterance unfolds. It is expected that when other information regarding the ambiguous referent is available, such as the Frequency and Dominance information available with Biased homophones, Sarcastic Prosody may not receive the same priority as an informational cue compared to when the information associated with the homophone is lacking.

Both the Graded Salience Hypothesis and Relevance Theory, as introduced in Chapter 2, argue that information provided in the utterance contributes to how a listener

will interpret the utterance, such that different combinations of information will result in different interpretations (Giora, 1997; Sperber & Wilson, 1995). In particular, if the sarcasm is taken to indicate marked, important information, as was the argument presented in section 2.3, the theories argue that this additional information is useful given particular contexts, and predictions can be generated from each viewpoint. The predictions of Graded Salience and Relevance Theory, in addition to the interpretations that can be offered by Direct Access View for the processing of sarcasm, are further detailed below (Gibbs, 1986; Gibbs, 2002).

Predictions

1. Direct Access View

An early account of sarcasm processing presented and further developed by Gibbs (1986; Gibbs, 2002) suggests that under the conditions offered within the current experiment (sufficient, realistic social contexts that the sarcasm is placed within), a sarcastic interpretation of a discourse should not depend on the failure of a sincere interpretation. If this is the case, then processing across all experimental items should be similar in both conditions, such that responses to items containing Sarcasm should have similar response times in making the judgment (choosing a context), when compared to items heard in the Sincere condition. If there are differences in decision times, then this approach is not supported by the current work[1].

---

[1] The other perspectives discussed in the theoretical review, such as Echoic Reminder Theory and the Muting Hypothesis (Kreuz & Glucksberg, 1989, Dews & Winner, 1995), may also suggest that given contexts that are more geared toward one interpretation via cues such as social contextual information, sarcasm will be easier to interpret, than in situations where this information is lacking. However, given that these theories tend to

2. Graded Salience Hypothesis

Overall, using the Graded Salience Hypothesis framework (Giora, 1997), sarcasm is treated as a cue that has the possibility of making an interpretation more salient, when combined with previous cues (such as Dominance and Frequency). This salience cue can be added to ongoing competing discourse interpretations. Using this framework, and continuing from the Linking Hypothesis, it is likely that Sarcasm will interact with the previous cues, such as to contribute to the grading of the salience between actively competing interpretations. If it is interpreted as adding information and increasing the consideration of alternative meanings, sarcasm will be most useful in this framework when discourse interpretations are continuously competing until participants are forced to make a choice. Given this, there should also be differences when making a choice in these cases. If the choice *itself* does not differ between conditions, then at least the time to *make the choice* should differ (such that it takes longer to react in one condition versus the other, as Sarcastic Prosody has offered additional information). Given this framework, we can generate further predictions made using a Graded Salience Hypothesis framework for both Biased and Balanced homophones.

Biased Homophones: Graded Salience Hypothesis

When examining different interpretations, such as those competing in the case of a *Biased* homophone, there is already a difference in the likelihood of an interpretation of a particular meaning, given the Dominance information. If Sarcasm has an effect here, as measured by changes in fixation durations to the contextual depictions, in order for it to

---

apply to specific contextual situations, (e.g., echoing of social norms) predictions for the current paradigm are not offered beyond this note.

be considered useful under Graded Salience framework, it must increase fixations and consideration of the previously (e.g., at Sentence 1 baseline) least-considered alternative, and possibly decrease the salience of the other alternative. This also involves the Sarcasm having achieved status as a salient cue. Otherwise, giving the preceding information regarding Dominance and Frequency, we would not expect Sarcasm to impact resolution, as the other cues have already provided salience for an interpretation.

Balanced Homophones: Graded Salience Hypothesis

However, in the case of *Balanced* homophones, less information regarding the speaker's intended meaning may be available upon lexical access, as no Dominance information leads to a meaning favored for interpretation. In this case, when Sarcasm is utilized in a discourse following the presentation of a homophone, there should be differences in the amount that it increases fixation durations to both High and Low frequency alternative meanings (as Frequency is one of the few cues available), but it should provide an additional salience cue, that aids listeners in choosing an interpretation. Thus, here it may add Salience (consideration) to both alternatives, but it should do so in the form of a moderator.

Final Choice: Graded Salience Hypothesis

When examining final interpretation of the discourse for Biased and Balanced homophones, when hearing Sarcastic Prosody participants should be more likely to assign a "sarcastic" interpretation of the discourse as referring to an initially less likely alternative. The interpretation that is "less likely" can be identified only by examining the choice in the Sincere Prosody condition, and using it for comparison. This is given that Sarcasm was utilized as a salient cue. Therefore, there should be a change in the

choice proportions (given Sarcastic Prosody) such that the most likely interpretations given Sincere Prosody, decrease. Additionally, Sarcasm should have an effect on how long it takes participants to make a final choice, such that they consider an alternative, and therefore demonstrate different reaction times (a main effect) in the Sarcastic Prosody condition.

3. Relevance Theory Framework

In general, Relevance Theory framework would also predict that the informational cues of Dominance, Frequency, and Sarcasm interact when participants choose an interpretation of the meaning of a homophone, and therefore discourse (Sperber & Wilson, 1995). If one discourse interpretation is already salient, in order to change said interpretation, Sarcasm must serve as a powerful contextual effect, and be seen as relevant information. If this is not the case, the interpretation of the discourse should not reflect a change, but perhaps the interpretation already chosen will be strengthened. Following a Relevance Theory approach, it is expected that this will lead to some contexts in which sarcasm is beneficial, and serves as a contextual effect to strengthen a chosen interpretation, and others where it serves to highlight an alternative interpretation, thus, we would expect it to interact with a "change" variable (e.g., whether or not listener's continue fixating on the same context). Importantly, it should not serve both effects within the processing of a Biased or Balanced homophone in a given Social Contextual situation. Thus, it should not increase the likelihood of fixating on both contextual representations of the homophone (or have a main effect). Instead, the framework would predict an effect of Sarcasm interacting with characteristics of the homophones, such as Dominance and Frequency.

Biased Homophones: Relevance Theory

Contextual effects are hypothesized to work such that when an informational cue is seen as "relevant" it can affect the processing of the ongoing discourse interpretation. The contextual effects that Sarcasm was identified as being most likely to act as (identified in Chapter 2) were strengthening existing interpretations, or as a contradictory contextual effect. This distinction is considered dependent on the contribution of other information in the discourse when Sarcastic Prosody is introduced. In the Biased homophone condition, there is already a salient interpretation that listeners should have no reason to re-evaluate, given that the cues of Dominance and Frequency have provided a strong initial interpretation (as in the case of the Graded Salience Hypothesis). Thus, Sarcasm may not be considered relevant information, or strong enough to change the interpretation of the discourse when introduced following the homophone. In this case, Sarcasm should serve as a strengthening effect, (if it has an effect) for the already salient interpretation (at baseline). If however, Sarcasm *does* highlight an alternate interpretation, one could argue from a Relevance Theory perspective that it is serving as a contradictory contextual effect, and that the prior Dominance and Frequency were insufficient to override any effect of Sarcasm. Thus, it can readily (*a priori*) explain both outcomes.

Balanced Homophones: Relevance Theory

Turning to Balanced homophones, much less information is available for listeners. This provides a context in which Sarcasm is immediately considered more "relevant" and useful information. Relevance Theory would suggest that it would strengthen one meaning interpretation, at the expense of another. This highlighting, or

strengthening effect of a particular interpretation should interact with both Frequency and Social Contextual Ratings to a much greater extent in the case of a Balanced versus a Biased homophone, as listeners have fewer initial cues to work with. Thus, interaction of Sarcasm with other covariates in the Balanced condition is predicted from a Relevance Theory perspective because Sarcasm itself, gains relevance given the poverty of initial Dominance information.

Final Choice: Relevance Theory

Finally, the framework would suggest that within final choice selections, there should be differences only where sarcasm increases the processing of an alternative. Relevance theory would not see the absence of change in the Biased homophone condition as problematic, just that Sarcasm was not sufficiently relevant to serve as a contextual effect. Thus, we would expect the greatest effect of Sarcasm in the Balanced homophone conditions, interacting with the covariates.

Summary of Predictions

Thus, for these discourses Sarcasm should be serving to highlight information to differing extents according to the different theoretical perspectives, and the final interpretation of the discourse should be dependent on sarcasm's interaction with homophone Frequency and Bias information, as well Social Context. Both Graded Salience (Giora, 1997) as well as Relevance Theory (Sperber & Wilson, 1995) frameworks predict differences in interpretations based on interactions amongst the previous lexical information; described here as factors Frequency and Bias. Interactions are included within the results as postulated by theoretical guidance. Fixations at the

onset of the homophone are analyzed in order to observe baseline processing of the mention of the homophone.

7.5 Results For Experiment 2

First, accuracy in choosing the correct picture was calculated for the filler items as a measure of participant's investment in the experiment. Accuracy ranged from 93-100%, and no participants were excluded due to poor accuracy. As previously noted within the general procedures portion of the common Methods section (section 7.2), data was prepared for analysis by first creating fixation reports from raw eye position data. This resulted in a loss of < 2% of the trials. SR Research's DataViewer (SR Research Ltd., 2013) was used for this process, and the data was then merged with item-specific covariate information (as well as the data collected in Experiment 1) using R software. Fixations that were less than 60 *ms* and greater than 2000 *ms* were removed from the data, as they were considered outliers. Mixed-effects multi-level models were used to address the predictions for the theoretical perspectives (see section 7.4). Data from models of eye-movements generated according to the Linking Hypothesis above, as well as behavioral measures, is analyzed.

7.5.1 Eye-Movement Models

Following the Linking Hypothesis established, the analysis of eye-movement data focused on the two specific linguistic events identified in the Linking Hypothesis: the onset of the homophone in Sentence 1, and the presence or absence of Sarcastic Prosody in Sentence 2. The Sentence 2 time frame was divided into 2000 *ms* sections to examine changes in processing over time. Additionally, given the different theoretical predictions

for the Biased (n = 20) and Balanced (n = 10) homophone pairs, they were analyzed separately.

Sentence 1

For homophones with differences in both Dominance and Frequency classifications (Biased homophones), there should be differences in fixation durations based on the information when the homophone is mentioned, or at baseline. Similarly with Balanced homophones, but hypothesized differences would only be generated by Frequency. A window to model fixation durations at this point was generated by restricting fixations to those that began after the onset of the homophone during the sentence, and ended no more than 750 *ms* after. Thus, no fixations that began before the homophone was begun were included within the analysis. Social Contextual Ratings and Picture Proportions (from Experiment 1) were mean-centered for all analyses (in the study) and used as covariates for the factors Frequency and Dominance (within the Biased homophones, Frequency only within the Balanced homophones). Results are reported for the full 750 *ms* time window as well as first fixations during the window only.

Biased Homophones

For the full 750 *ms* time window, within the fixation durations there was a significant Frequency by Bias interaction (such that for the Subordinate meaning, fixations to High Frequency items were significantly longer), and a main effect of Bias that approached significance (such that Subordinate meanings were fixated on longer)

(see Table 7.2, Model 1). The means can be viewed in Figure 7.1. This analysis was then repeated, and restricted to first fixations. With the restriction, the main effect of Dominance became significant within the model, and a main effect of Frequency approached significance. Their interaction remained significant (see Table 7.2, Model 2). The means can be viewed in Figure 7.2. As an illustration, the model being tested can be viewed in [5]. Additional models for other analyses have modifications of predictors based off this model.

[5] $Y_{ij} = \beta_{0i} + \beta_1$*PictureProportion$_{1ij} + \beta_2$*SocialContextRating$_{2ij} + \beta_3$*Frequency$_{3ij} + \beta_4$*Dominance$_{4ij} +$ β_5*Frequency*Dominance$_{5ij} + b_{i1}$* Subject$_{1i} + b_{i2}$ * Item$_{2j} + \varepsilon_{ij}$

Table 7.2. Models of fixations to Biased homophones in Sentence 1.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, Frequency, and Bias. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes High Frequency, Dominant as reference groups. Significant estimates are bolded.

Model 1. *Biased Homophones full 750 ms time window*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	228.02	8.96	25.44	**0.001**
Picture Proportions	6.03	61.06	0.10	= 0.92
Social Contextual Rating	1.21	6.26	0.19	= 0.85
Low Frequency	10.68	9.50	1.12	= 0.26
Subordinate Meaning	16.41	10.66	1.54	= 0.12
Low Frequency*Subordinate	-26.81	12.79	-2.10	**0.04**

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	798.91	28.27
Item	(Intercept)	34.64	5.89
Residual		7449.12	86.31

Model Fit Measures	
AIC	10,162
BIC	10,204
logLik	-5072

Model 2. *Biased Homophones first fixation after homophone onset only*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <

Intercept	235.72	11.24	20.97	**0.001**
Picture Proportions	3.98	83.11	0.05	= 0.96
Social Contextual Rating	3.14	8.81	0.35	= 0.72
Low Frequency	21.84	13.25	1.65	**0.10**
Subordinate Meaning	31.94	15.23	2.10	**0.04**
Low Frequency*Subordinate	-46.31	18.02	-2.57	**0.01**

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	652.71	25.55
Item	(Intercept)	47.21	6.87
Residual		8805.73	93.84

Model Fit Measures	
AIC	6148
BIC	6186
logLik	-3065

Combined, the data indicate that for Biased homophones, both Frequency and Dominance have early effects on directing eye movements at the onset of the word in speech, such that High Frequency-Subordinate meanings, when pictured, draw additional processing. Additionally, the models demonstrate that any effects of the pictures themselves have dissipated (no significant effects of Picture Proportions) and that Social Context, if it is considered in the task, does not have an immediate effect.

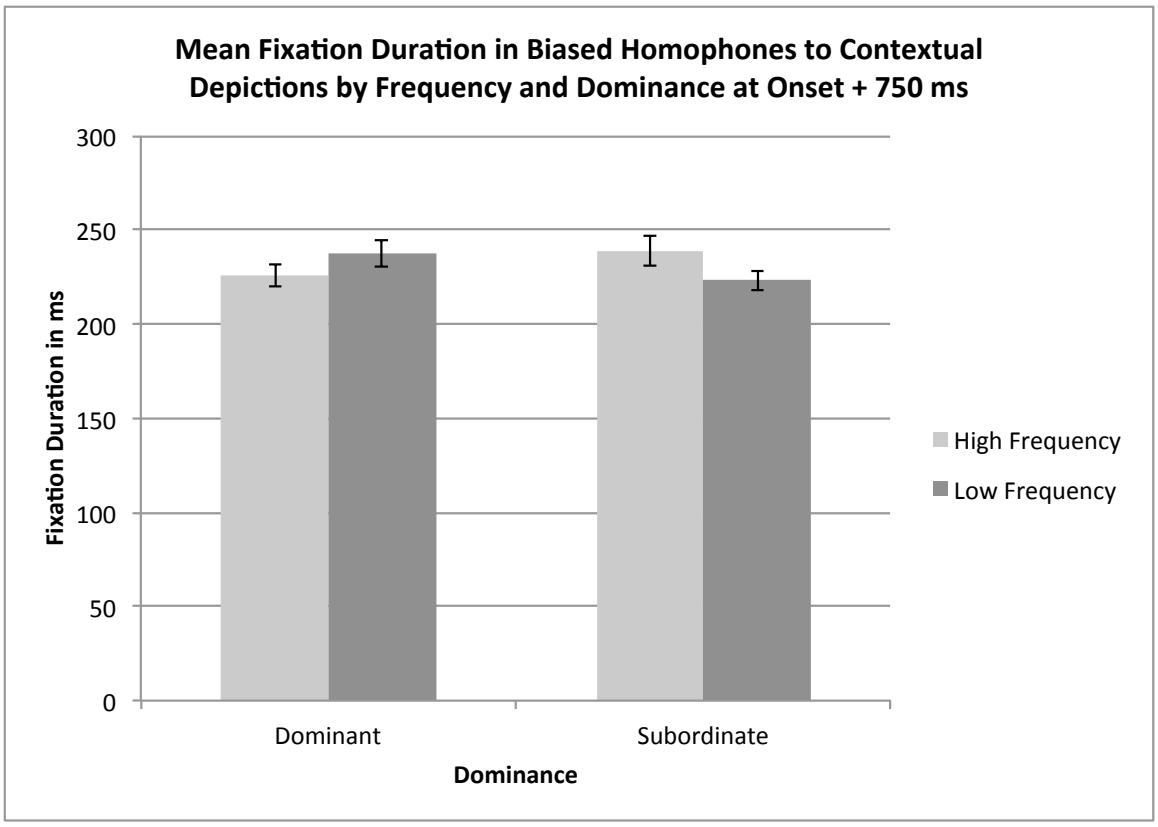

Figure 7.1. Mean Sentence 1 fixation durations in *ms* at homophone onset to Biased homophones by Dominance and Frequency. Standard error of mean depicted.

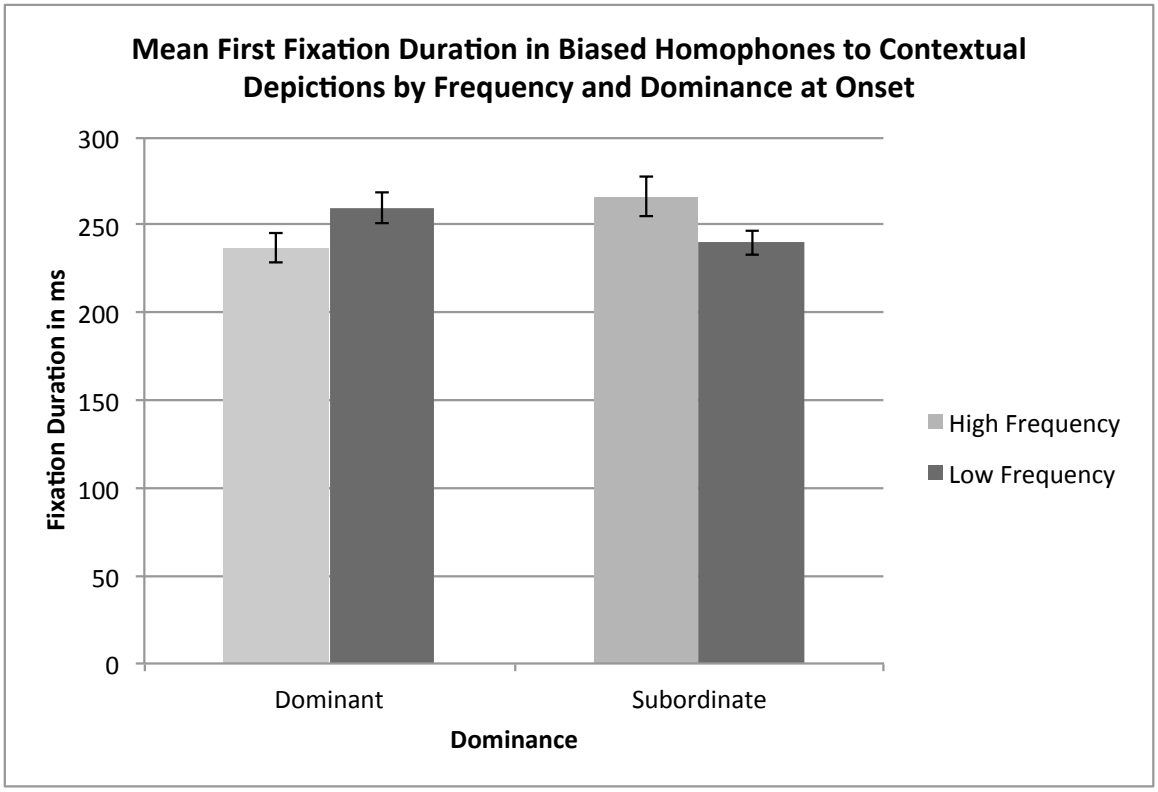

Figure 7.2. Mean Sentence 1 first fixation durations in *ms* at homophone onset to Biased homophones by Dominance and Frequency. Standard error of mean depicted.

Balanced Homophones

When similar models were used to analyze the Balanced homophones (see [6]), if Picture Proportions and Social Contextual Ratings were used in the model as covariates, it did not converge. However, when they were removed, a model of all fixations during the 750 *ms* time window after homophone onset did converge, but the effect of Frequency only approached significance (such that Low Frequency contextual representations were fixated on *slightly* longer initially) (see Table 7.3, Model 1). Conversely, if only the first fixation durations were analyzed, the model converged with Picture Proportions and Social Contextual Ratings in it, but neither the two nor

Frequency were significant predictors of first fixation duration (see Table 7.3, Model 2). This indicates that given a Balanced homophone, there is no readily available factor to distribute attention favorably to one meaning or another, resulting in equal attention to the contexts being observed. In this time period, it appears that not even Frequency classification has the ability to direct attention. This predictor will be analyzed once more to determine if it interacts with Sarcastic Prosody in Sentence 2, as was predicted by both Graded Salience and Relevance Theory framework.

[6] $Y_{ij} = \beta_{0i} + \beta_1 * \text{PictureProportion}_{1ij} + \beta_2 * \text{SocialContextRating}_{2ij} + \beta_3 * \text{Frequency}_{3ij} + b_{i1} * \text{Subject}_{1i} + b_{i2} * \text{Item}_{2j} + \varepsilon_{ij}$

Table 7.3. Models of fixations to Balanced homophones in Sentence 1.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, and Frequency. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes High Frequency reference group. Significant estimates are bolded.

Model 1. *Balanced Homophones full 750 ms time window*

Fixed Effect Coefficient Estimates		β *Est.*	*Std. Error*	*t*	*p* <
Intercept		217.87	8.08	26.97	**0.001**
Low Frequency		12.16	8.13	1.50	= 0.13
Random Effects	*Name*	*Variance*	*Std. Dev.*		
Subject	(Intercept)	735.63	27.117		
Item	(Intercept)	0.00	0.00		
Residual		6969.47	83.48		
Model Fit Measures					
AIC	5436				
BIC	5457				
logLik	-2713				

Model 2. *Balanced Homophones first fixation after homophone onset only*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	230.83	12.02	19.21	**0.001**
Picture Proportions	-73.78	94.11	-0.78	= 0.43
Social Contextual Rating	3.69	12.68	0.29	= 0.77

Low Frequency		18.15	14.56	1.25	= 0.21

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	288.75	16.99
Item	(Intercept)	0.00	0.00
Residual		8725.02	93.41

Model Fit Measures	
AIC	3132
BIC	3157
logLik	-1559

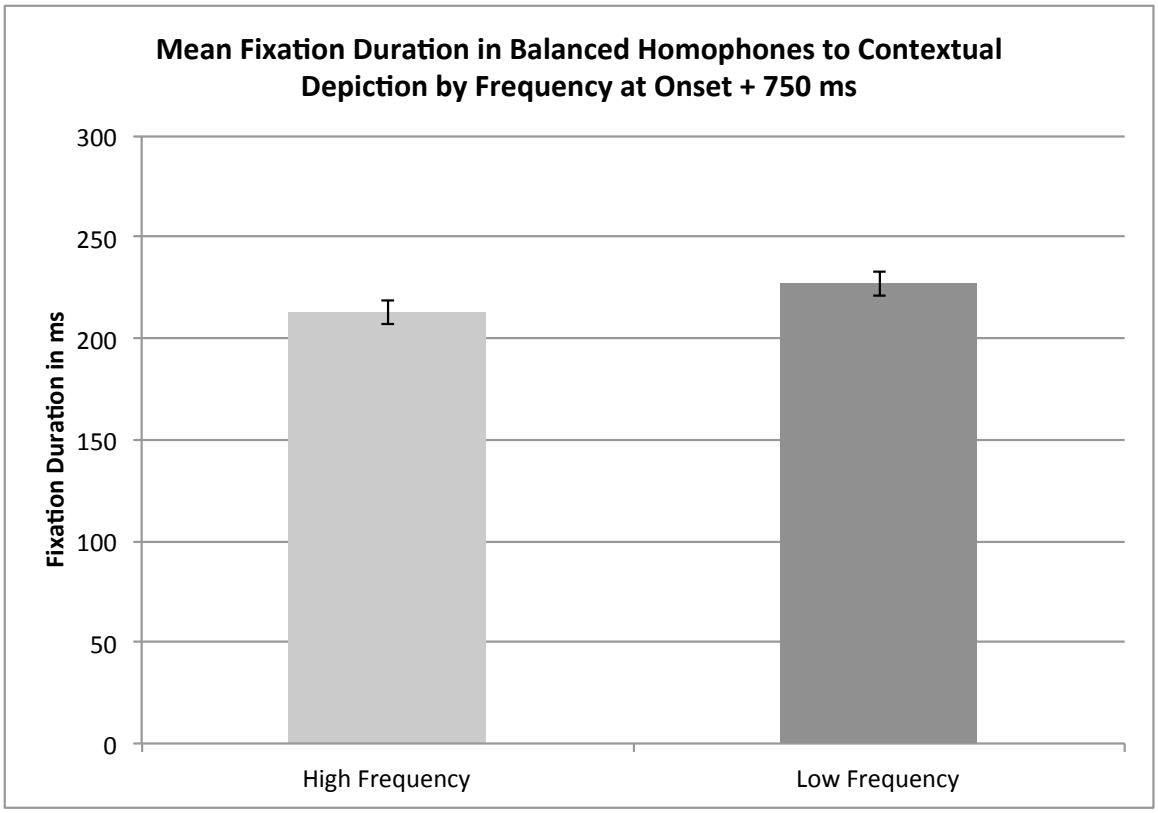

Figure 7.3. Mean Sentence 1 fixation durations in *ms* at homophone onset to Balanced homophones by Frequency. Standard error of mean depicted.

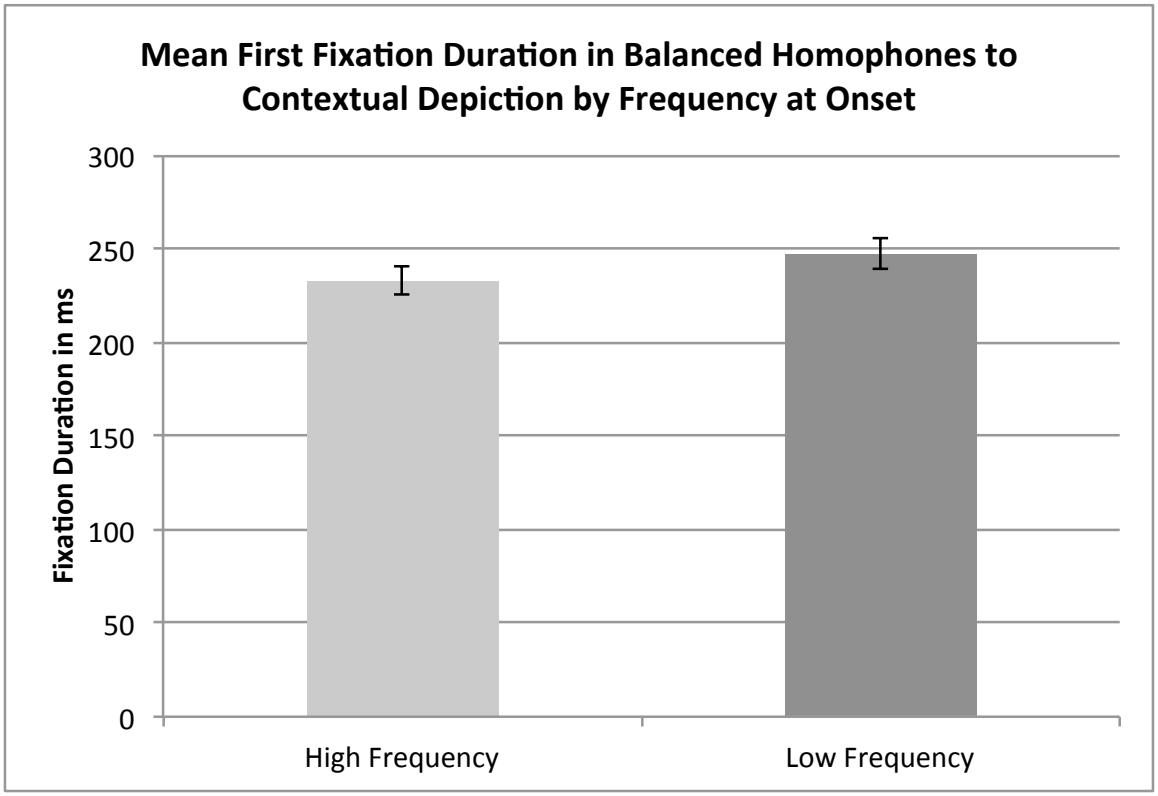

Figure 7.4. Mean Sentence 1 first fixation durations in *ms* at homophone onset to Balanced homophones by Frequency. Standard error of mean depicted.

Sentence 2

Once again, the data were analyzed by modeling fixations according to the Linking Hypothesis. Sentence 2 models were broken into the first 2000 *ms* and the second 2000 *ms* of the sentence in order to analyze the early and late effects of Sarcastic Prosody on processing. Sarcasm began within the first 2000 *ms* of the sentence, but varied by item as to the exact location of introduction. Additionally, since one of the components of the Sarcastic Prosody manipulation was lengthening the item, time-locking to the "beginning" of a sarcastic statement altered the entire sentence.

Biased Homophones

For Biased homophones, during the first 2000 *ms* of Sentence 2, since Sarcastic Prosody was becoming available, the predictors tested included results of the participants' Faux Pas task performance, the Social Contextual Rating for the item, Frequency, and Dominance information (see [7]). A final variable that was introduced was whether the participant changed to fixate on the opposite picture next, or not.

[7] $Y_{ij} = \beta_{0i} + \beta_1*\text{FauxPas}_{1ij} + \beta_2\text{Frequency}_{2ij} + \beta_3*\text{Dominance}_{3ij} + \beta_4*\text{SocialContextRating}_{4ij} + \beta_5*\text{Sarcasm}_{5ij} + \beta_6*\text{Change}_{6ij} + \beta_7*\text{Frequency}*\text{Dominance}_{7ij} + \beta_8*\text{SocialContextRating}*\text{Sarcasm}_{8ij} + \beta_9*\text{SocialContextRating}*\text{Change}_{9ij} + \beta_{10}*\text{Sarcasm}*\text{Change}_{10ij} + \beta_{11}*\text{SocialContextRating}*\text{Sarcasm}*\text{Change}_{11ij} + b_{i1}*\text{Subject}_{1i} + b_{i2}*\text{Item}_{2j} + \varepsilon_{ij}$

The model results can be viewed in Table 7.4, Model 1. Neither the covariates nor factor Sarcasm predicted fixation duration during the first half of the sentence, indicating high variability in how any given participant chose to fixate on a particular meaning.

Alternatively, within the second half of Sentence 2, there is a significant interaction between Frequency and Dominance ($p < 0.02$) (see Table 7.4, Model 2). Additionally, there is a main effect of Change, such that if a switch is made during the next fixation, the current fixation is longer. To understand these effects more completely, the means by Frequency and Dominance can be viewed in Figure 7.5. The interaction appears primarily driven by the difference in the Subordinate meaning condition of Bias, in which High Frequency items received longer fixation durations. This is a continuation from baseline at Sentence 1, suggesting that Sarcasm did not change final interpretation.

Table 7.4. Models of fixations to Biased Homophones in Sentence 2.

Data modeled according to predictors Faux Pas results, Frequency, Bias, Social Contextual Rating, Sarcastic Prosody and Change. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes High Frequency, Dominant, Sincere as reference groups. Significant estimates are bolded.

Model 1. *Biased Homophones first 2000 ms Sentence 2*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	370.98	74.33	4.99	**0.001**
Faux Pas	-80.30	75.47	-1.06	= 0.29
Low Frequency	0.54	8.97	0.06	= 0.95
Subordinate Meaning	-4.67	9.51	-0.49	= 0.62
Social Contextual Rating	-9.29	7.79	-1.19	= 0.23
Sarcastic Prosody	5.50	25.65	0.21	= 0.83
Change	35.29	34.04	1.04	= 0.30
Low Frequency*Subordinate	-12.06	11.00	-1.10	= 0.27
Social Rating*Sarcastic	-1.29	9.50	-0.14	= 0.89
Social Rating*Change	-7.40	12.51	-0.59	= 0.55
Sarcastic Prosody*Change	25.94	50.17	0.52	= 0.61
Social Rating*Sarcastic*Change	-11.78	11.29	-0.64	= 0.52

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	1497.50	38.70
Item	(Intercept)	211.64	14.55
Residual		15.289.30	123.65

Model Fit Measures	
AIC	36,079
BIC	36,169
logLik	-18,025

Model 2. *Biased Homophones second 2000 ms Sentence 2*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	330.93	111.13	2.98	**0.003**
Faux Pas	-28.50	118.43	-0.18	= 0.85
Low Frequency	-2.03	17.62	0.06	= 0.95
Subordinate Meaning	22.29	19.87	1.13	= 0.26
Social Contextual Rating	-8.57	14.79	-0.52	= 0.60
Sarcastic Prosody	-5.99	10.56	-0.77	= 0.44
Change	38.56	14.90	2.29	**0.02**
Low Frequency*Subordinate	-54.90	22.42	-2.41	**0.02**
Social Rating*Sarcastic	-2.59	17.64	-0.41	= 0.68
Social Rating*Change	-42.54	25.26	-1.32	= 0.19

Sarcastic Prosody*Change	-19.78	21.71	-0.58	= 0.56
Social Rating*Sarcastic*Change	44.92	36.18	1.16	= 0.24

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	3461.25	58.83
Item	(Intercept)	495.23	22.25
Residual		25,477.61	159.62

Model Fit Measures	
AIC	16,331
BIC	16,408
logLik	-8150

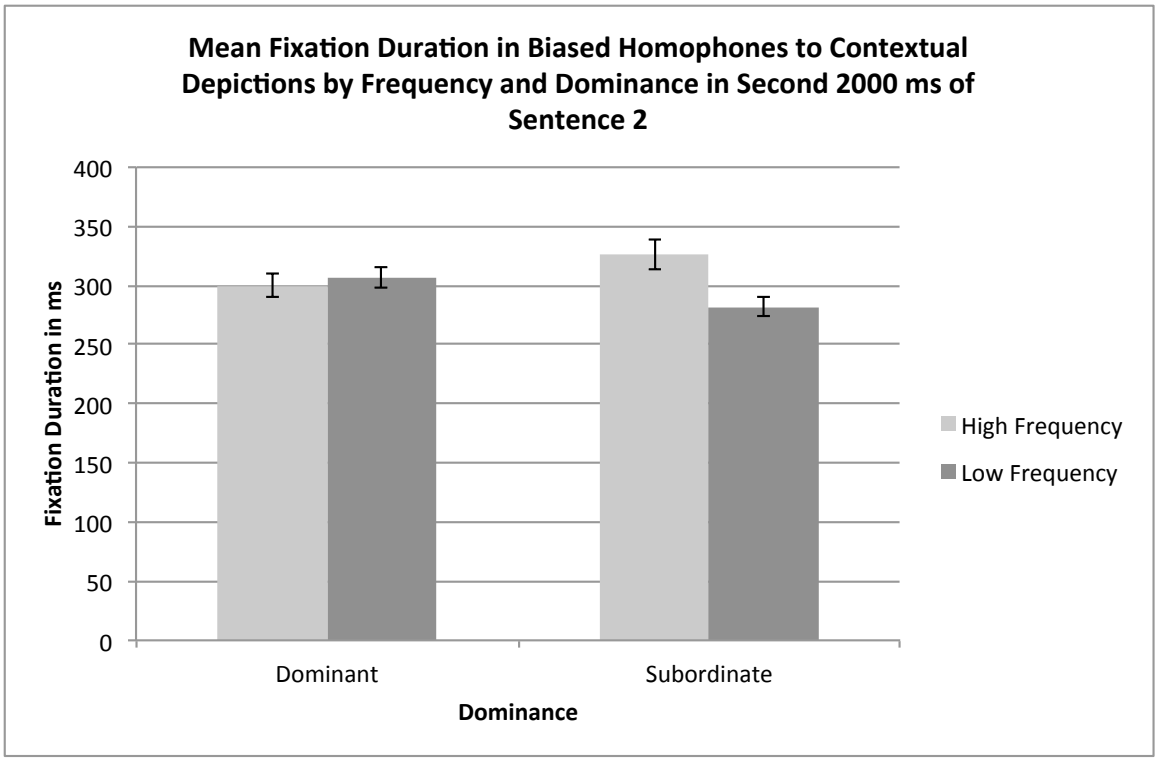

Figure 7.5. Mean Sentence 2 fixation durations by conditions Frequency and Dominance.

Balanced Homophones

Finally, fixations to the Balanced homophones by condition were also considered for Sentence 2. The same covariates were used as within the Biased homophone models, with the exception of Bias as it did not exist as a predictor for this condition (see [8]).

[8] $Y_{ij} = \beta_{0i} + \beta_1$*FauxPas$_{1ij}$ + β_2Frequency$_{2ij}$ + β_3*SocialContextRating$_{3ij}$ + β_4*Sarcasm$_{4ij}$ + β_5*Change$_{5ij}$ + β_6*SocialContextRating*Sarcasm$_{6ij}$ + β_7*SocialContextRating*Change$_{7ij}$ + β_8*Sarcasm*Change$_{8ij}$ + β_9*SocialContextRating*Sarcasm*Change$_{9ij}$ + b_{i1}* Subject$_{1i}$ + b_{i2} * Item$_{2j}$ + ε_{ij}

As opposed to the Biased homophones, there were significant effects of the predictors and covariates within the first 2000 *ms* of Sentence 2. There was a main effect of Change, and also an interaction of Sarcastic Prosody with Social Contextual Rating and of Sarcastic Prosody with Change (see Table 7.5, Model 1). The interaction of Social Contextual Rating and Sarcastic Prosody was such that if the sentence contained Sarcastic Prosody, higher Social Contextual Ratings contributed to an increase in fixation time. The means of Sarcastic Prosody by the Change variable are presented in Figure 7.6. As there was not a significant effect of Frequency classification in Sentence 1 in the Balanced condition, it is likely that listeners turned to evaluation of the Social Context of the conversation and begin using the information as it became available, to complete the task at hand. This in turn elevated the importance of the Sarcastic Prosody, as predicted by both Graded Salience and Relevance Theory (Giora, 1997; Sperber & Wilson, 1995).

In the model of processing for the second 2000 *ms* of Sentence 2, the interaction of Social Contextual Ratings and Sarcastic Prosody remained significant in the same direction as in the first half of Sentence 2 (longer fixations given higher Social Contextual Ratings). Additionally, both Social Contextual Ratings and Sarcastic Prosody achieved moderately significant main effects ($p = 0.10$), both of which involved *reduced*

fixation durations given an increase in Social Context Ratings, or the presence of Sarcastic Prosody (see Table 7.5, Model 2). However, interpretation of the interaction takes precedence. Finally, the effect of Social Contextual Ratings by Sarcasm and Change was also significant, such that if there was an upcoming Change, in the Sarcastic Prosody condition the fixation duration was shortened. When compared to the model for the second 2000 *ms* of Sentence 2 in the Biased condition, the data suggest listeners elevate the importance of the Sarcastic Prosody and Social Context given the lack of additional cues from the activation of the homophone. This provides useful information regarding the extent to which listeners prioritize information associated with ambiguous referents, and utilize sarcasm in assigning meaning to these ambiguities.

Table 7.5. Models of fixations to Balanced Homophones in Sentence 2.

Data modeled according to predictors Faux Pas results, Frequency, Social Contextual Rating, Sarcastic Prosody and Change. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes High Frequency, Sincere as reference groups. Significant estimates are bolded.

Model 1. *Balanced Homophones first 2000 ms Sentence 2*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	305.02	79.31	3.85	**0.001**
Faux Pas	-46.88	87.89	-0.55	= 0.58
Low Frequency	10.58	10.11	1.05	= 0.30
Social Contextual Rating	-14.67	12.32	-1.19	= 0.23
Sarcastic Prosody	-11.37	8.27	-1.38	= 0.17
Change	26.57	10.46	2.54	**0.02**
Social Rating*Sarcastic	30.28	14.14	2.14	**0.04**
Social Rating*Change	29.73	18.18	1.64	**0.10**
Sarcastic Prosody*Change	-7.03	15.42	-0.46	= 0.65
Social Rating*Sarcastic*Change	-37.32	26.24	-1.42	= 0.15

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	1747.06	41.80
Item	(Intercept)	192.86	13.89
Residual		14,388.08	119.95
Model Fit Measures			
AIC	16,343		
BIC	16,411		
logLik	-8159		

Model 2. *Balanced Homophones second 2000 ms Sentence 2*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	325.96	101.10	3.22	**0.002**
Faux Pas	-37.50	108.34	-0.35	= 0.73
Low Frequency	13.86	11.39	1.22	= 0.22
Social Contextual Rating	-23.78	14.52	-1.64	**0.10**
Sarcastic Prosody	-15.91	9.70	-1.64	**0.10**
Change	18.20	12.22	1.49	= 0.14
Social Rating*Sarcastic	44.91	16.77	2.68	**0.008**
Social Rating*Change	30.10	20.55	1.46	= 0.14
Sarcastic Prosody*Change	11.47	17.99	0.64	= 0.52
Social Rating*Sarcastic*Change	-78.19	29.75	-2.63	**0.009**

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	2949.80	54.31
Item	(Intercept)	293.53	17.13
Residual		30,810.78	175.53
Model Fit Measures			
AIC	27,109		
BIC	27,183		
logLik	-13,542		

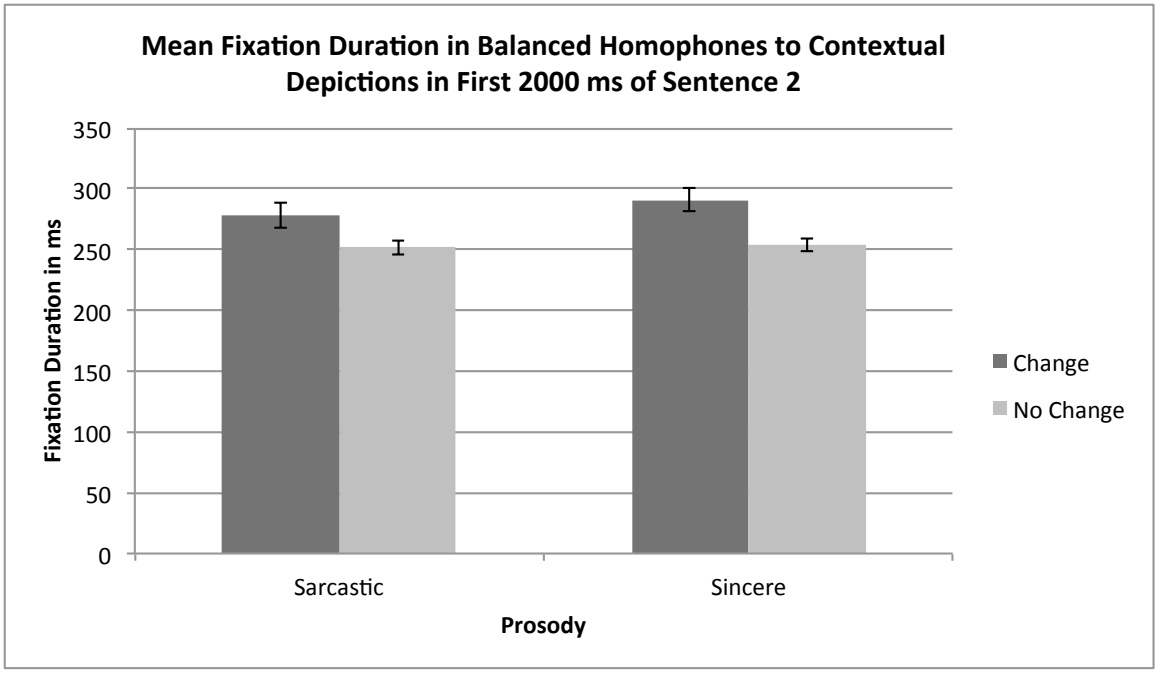

Figure 7.6. Mean Sentence 2 fixation durations for first 2000 *ms*. Fixations to Balanced homophones by Sarcasm and Change. Standard error of mean depicted.

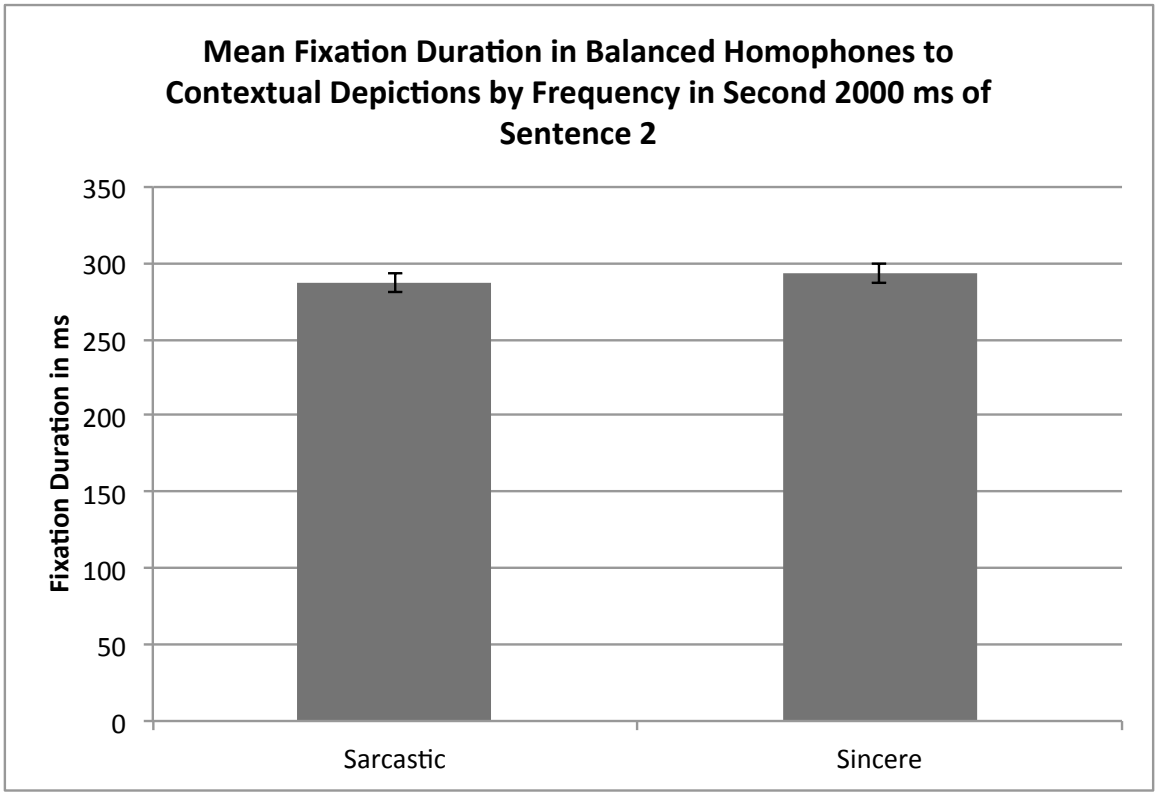

Figure 7.7. Mean Sentence 2 fixation durations for second 2000 *ms*. Fixations to Balanced homophones by Sarcasm. Standard error of mean depicted.

Summary

It appears that the importance of Sarcastic Prosody when following an ambiguous reference is moderated as predicted by the additional lexical information that is associated with the homophone. When homophones were examined by condition of Bias (Dominant versus Subordinate versus Balanced), Sarcastic Prosody affected the processing of the discourse differently. For example, in Sentence 1 at the offset of the homophones, in the Biased condition, Frequency and Dominance information were immediately activated, with longer fixations to a Subordinate, High Frequency meaning (see Table 7.2). In the same time period, in the case of the Balanced homophones,

Frequency information was not. However, within Sentence 2, while initially (the first 2000 *ms*) none of the predictors were significant in contributing to the fixation durations in the Biased conditions (see Table 7.4, Model 1), Social Contextual Ratings and Sarcasm interacted immediately during the same time window for the Balanced homophones, such that given Sarcastic Prosody, meanings with higher Social Contextual Ratings received more fixations (see Table 7.5, Model 1). Additionally, during the second half of Sentence 2, within the Biased homophone model Frequency and Dominance information was once again significant, while Sarcastic Prosody was never a significant factor. Thus, even later in processing, listeners continued paying more attention to the lexical level information. However, in the case of the Balanced homophones, downstream in processing, the interaction of Sarcastic Prosody and Social Context *strengthened*. To determine if these effects persist through to the final choice, behavioral data was examined.

As a final note, additional models were tested for Biased homophones, which included Sarcastic Prosody tested specifically as interacting with the Dominance and Frequency variables. However, the interaction of Sarcastic Prosody was not significant within these models and due to the heightened complexity of the models, they are not reported within these analyses so as to remain as tractable as possible. Sarcasm was also tested as interacting with Frequency in Balanced homophones, and the same null effects were observed. Thus, the results reported here examine the effects of Sarcasm and Social Context further, with the interaction of Sarcasm and Social Context, and Dominance and Frequency information, as independent pairs within the models.

7.5.2 Behavioral Data

Reaction Times

In order to test processing times, reaction times were collected for responses to the question given, as noted in the common methods. Reaction times were grouped by the contextual depiction chosen. Reaction times greater than 4000 *ms* were immediately removed from the analysis as they were considered outliers (4 responses).

Biased Homophones

First, a multilevel mixed-effects model with random intercepts for subjects and items was used to test for differences in reaction times for the Biased homophones. Within the model, Frequency, Dominance, Social Contextual Rating and Sarcasm served as predictors. Within the first model containing a Social Contextual Ratings predictor, the Frequency by Dominance interaction was moderately significant ($p = 0.07$) (see Table 7.6, Model 1). When the Social Contextual Ratings variable was removed, model fit improved and Dominance and the Frequency by Dominance interaction became significant, such that it took longer to respond with a choice of High Frequency, Dominant meaning, than a High Frequency, Subordinate meaning (p's < 0.05) (see Table 7.6, Model 2). While Sarcasm was not a significant factor in the model, the means per condition divided by Sarcasm can be viewed in Figure 7.8, while the means collapsed can be viewed in Figure 7.9. This suggests that in the case of Biased homophones, Sarcastic Prosody had a negligible effect on reaction time compared to the Frequency and Bias variables. Thus, Sarcasm appears to have had little to no effect for Biased homophones in the given the context of the conversation.

Balanced Homophones

Next, similar models were used to test reaction time in making a final choice when considering the Balanced homophones (Table 7.6, Model 3), with covariates Social Contextual Rating and Frequency. Here, Social Contextual Ratings were a significant predictor, and Sarcasm interacted with Frequency to determine reaction times. Additionally, Frequency and Sarcasm interacted to predict reaction times and Frequency became significant when Social Contextual Ratings were removed from the reaction time models, such that Sarcasm reduced reaction times to High Frequency selections, and lengthened reaction times to Low Frequency selections, compared to Sincere prosody (Table 7.6, Model 4). Thus, it appears that within the reaction time data for Balanced homophones, Sarcastic Prosody does finally interact with Frequency information. These effects can be viewed within Figure 7.10. The tested model is illustrated in [9] for Biased homophones, and [10] for Balanced.

[9] $Y_{ij} = \beta_{0i} + \beta_1$*SocialContextRating$_{1ij} + \beta_2$Frequency$_{2ij} + \beta_3$*Dominance$_{3ij} + \beta_4$*Sarcasm$_{4ij} +$ β_5*Frequency*Dominance$_{5ij} + \beta_6$*Frequency*Sarcasm$_{6ij} + \beta_7$*Dominance*Sarcasm$_{7ij} +$ β_8*Frequency*Dominance*Sarcasm$_{8ij} + b_{i1}$* Subject$_{1i} + b_{i2}$ * Item$_{2j} + \varepsilon_{ij}$

[10] $Y_{ij} = \beta_{0i} + \beta_1$*SocialContextRating$_{1ij} + \beta_2$Frequency$_{2ij} + \beta_3$*Sarcasm$_{3ij} + \beta_4$*Frequency*Sarcasm$_{4ij} +$ b_{i1}* Subject$_{1i} + b_{i2}$ * Item$_{2j} + \varepsilon_{ij}$

Table 7.6. Models of Experiment 2 reaction time data by Bias.

Data modeled by Bias according to predictors 1) Sarcastic Prosody and Social Contextual Rating 2) Sarcastic Prosody, with covariates Frequency and Dominance of homophone where applicable. Models are of reaction times, with subjects and items containing random intercepts. For Biased conditions, intercept estimate includes High Frequency, Dominant, Sincere condition as reference groups, while Balanced intercepts do not contain Bias. Significant effects are bolded.

Model 1. *Biased Homophones Sarcastic Prosody & Social Contextual Ratings*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	559.00	220.51	2.54	**0.02**
Social Contextual Rating	96.05	62.63	1.53	= 0.13
Low Frequency	-67.25	112.29	-0.60	= 0.55
Subordinate Meaning	-168.50	129.93	-1.30	= 0.19
Sarcastic Prosody	110.30	107.68	1.02	= 0.31
Low Frequency*Subordinate	278.92	152.83	1.83	**0.07**
Low Frequency*Sarcastic Prosody	-16.97	144.28	-0.12	= 0.91
Subordinate*Sarcastic Prosody	-9.22	154.86	-0.06	= 0.95
Low Freq.*Subordinate*Sarcastic	-40.13	202.96	-0.20	= 0.84

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	88,964	298.27
Item	(Intercept)	12,760	112.96
Residual		3378,848	581.25

Model Fit Measures	
AIC	9330
BIC	9382
logLik	-4653

Model 2. *Biased Homophones Sarcastic Prosody Only*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	858.45	1.1.74	8.44	**0.001**
Low Frequency	-91.78	112.69	-0.81	= 0.42
Subordinate Meaning	-247.55	120.65	-2.05	**0.05**
Sarcastic Prosody	110.24	107.75	1.02	= 0.31
Low Frequency*Subordinate	310.72	153.10	2.03	**0.05**
Low Frequency*Sarcastic	-21.74	144.35	-.015	= 0.88
Subordinate*Sarcastic	-15.00	154.97	-0.10	= 0.92
Low Freq.*Subordinate*Sarcastic	-28.43	203.00	-0.14	= 0.89

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	87,719	296.17
Item	(Intercept)	14,570	120.70
Residual		338,072	581.44

Model Fit Measures	
AIC	9340
BIC	9388
logLik	-4659

Model 3. *Balanced Homophones Sarcastic Prosody & Social Contextual Ratings*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	932.76	122.56	7.61	**0.001**
Social Contextual Rating	190.35	107.71	1.77	**0.08**
Low Frequency	-143.23	128.23	-1.12	= 0.26
Sarcastic Prosody	-175.64	140.69	-1.25	= 0.21
Low Frequency*Sarcasm	323.29	173.79	1.86	**0.06**

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	137,857.3	371.29
Item	(Intercept)	2503.9	50.04
Residual		456,185.3	675.42

Model Fit Measures	
AIC	4759
BIC	4789
logLik	-2372

Model 4. *Balanced Homophones Sarcastic Prosody Only*

Fixed Effect Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	1018.07	118.47	8.59	**0.001**
Low Frequency	-250.30	121.18	-2.07	**0.04**
Sarcastic Prosody	-158.32	140.45	-1.13	= 0.25
Low Frequency*Sarcasm	314.57	173.86	1.81	**0.07**

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	137,984.6	371.46
Item	(Intercept)	7212.1	84.92
Residual		456,012.9	675.29

Model Fit Measures	
AIC	4771
BIC	4797
logLik	-2379

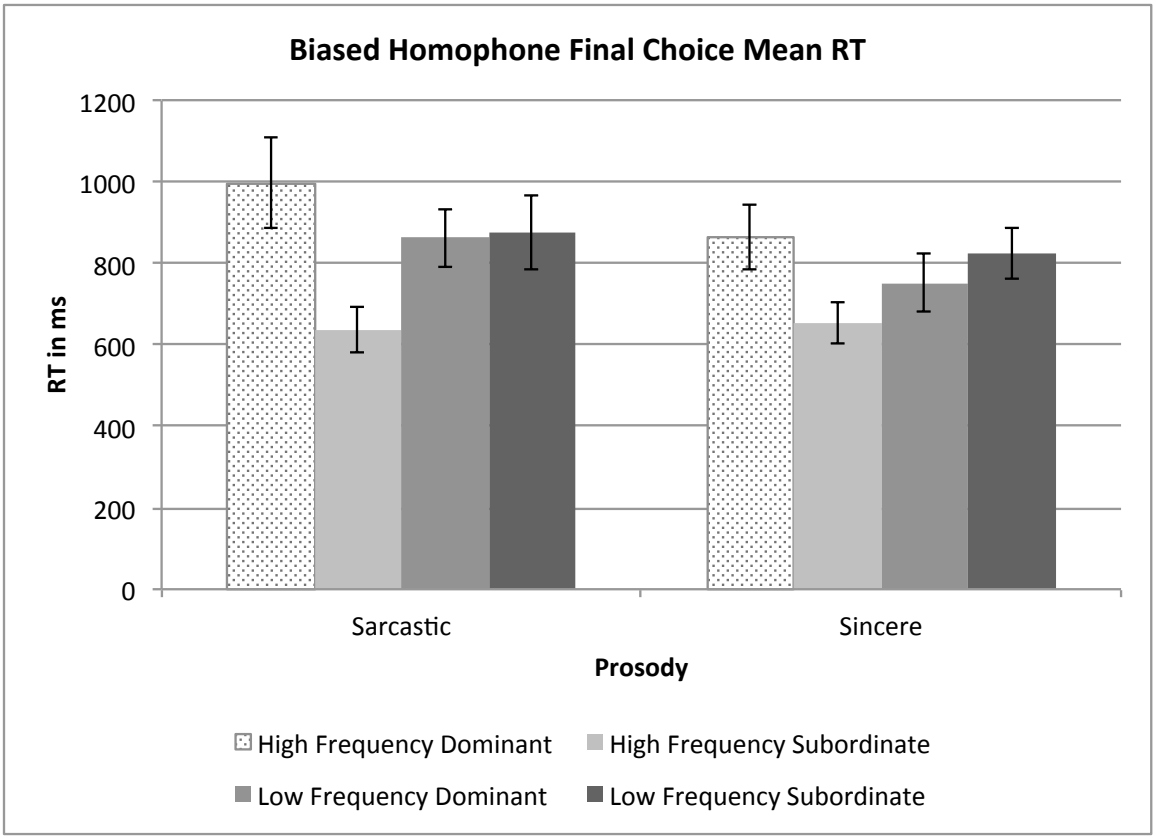

Figure 7.8. Mean reaction times in *ms* for responses to Biased homophones in the Sincere versus Sarcastic Prosody conditions of Experiment 2, according to Frequency and Dominance. Sarcasm factor not significant. Standard error of mean depicted.

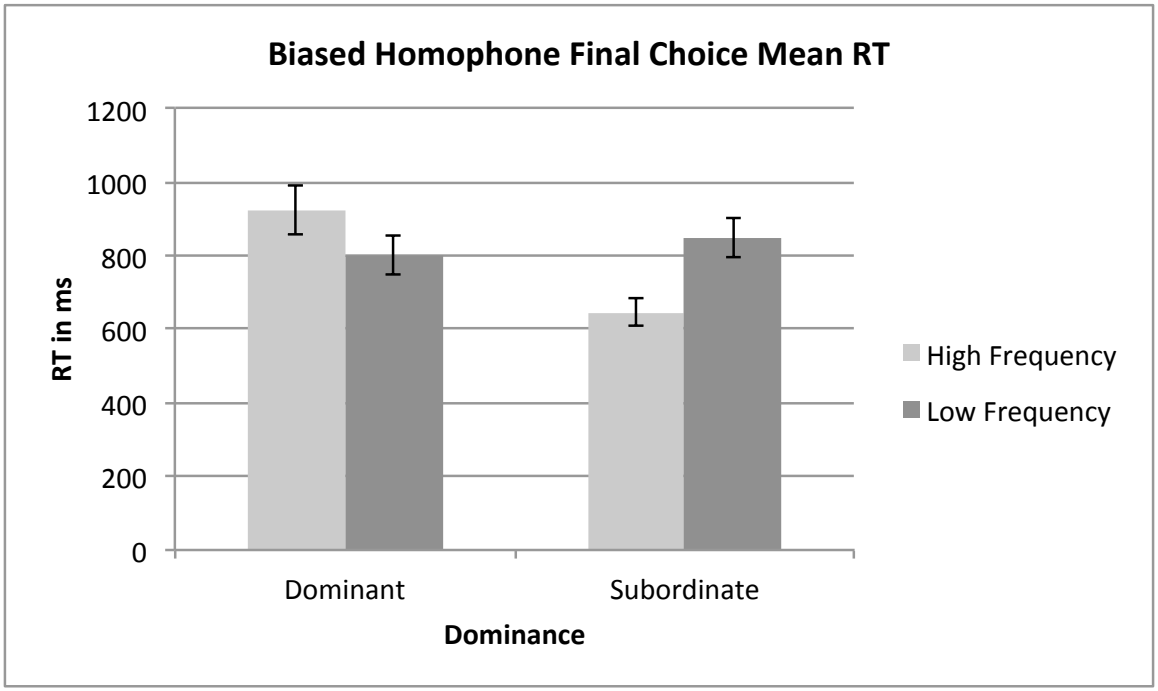

Figure 7.9. Mean reaction times in *ms* for responses to Biased homophones according to Frequency and Dominance. Significant effect of Dominance ($p < 0.05$), and Frequency by Dominance interaction, such that in the Subordinate condition, the difference between High and Low Frequency Response times was different ($p < 0.05$). Standard error of mean depicted.

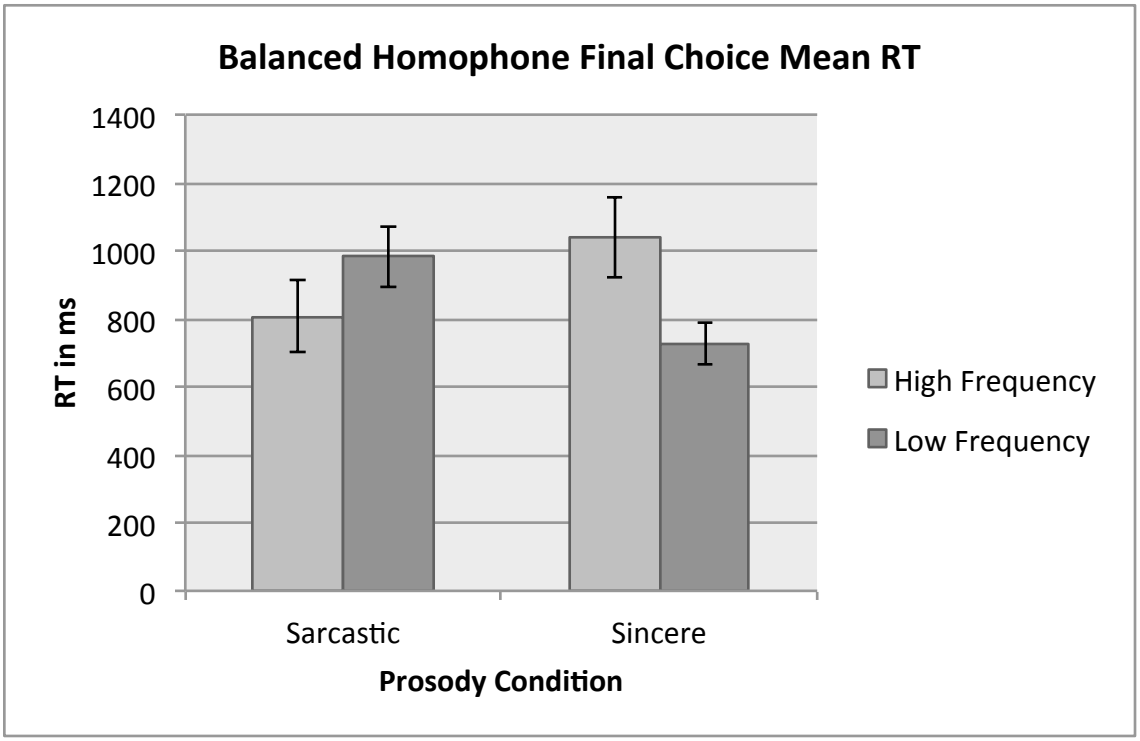

Figure 7.10. Mean reaction times in *ms* for responses to Balanced homophones in the Sincere versus Sarcastic Prosody conditions of Experiment 2, according to frequency, significance at .05. Standard error of mean depicted.

Final Choice Proportions

In addition to the reaction time data, items were analyzed by condition to determine the proportion of responses to each of the contextual depictions, given sarcastic versus sincere prosody. Chosen responses by condition were analyzed (answers were quantified into six conditions of possible context descriptions dependent on a combination of Dominance and Frequency, which can be viewed in Table 7.7) and a Chi-Square test for Independence was completed on the resulting proportions. The test found no differences by Sarcasm, thus the effects on final choice of Sarcasm appear solely in the reaction time data. Thus, while it appeared to affect the processing of different

possible homophone interpretations, sarcasm did not directly affect resolution, as measured by differences in the forced-choice task.

Table 7.7. Experiment 2 final choice proportions by factor Sarcasm.

Contextual Depiction Chosen	Sincere	Sarcastic
Balanced High Frequency	12.2%	10.75%
Balanced Low Frequency	21.4%	22.25%
Biased Subordinate Low Frequency	21.6%	22.25%
Biased Subordinate High Frequency	13.2%	14%
Biased Dominant Low Frequency	18.2%	17%
Biased Dominant High Frequency	13.4%	13.75%

In summary, the effects of Sarcastic Prosody on resolving an ambiguous reference when it follows appear to be moderated by characteristics of the homophone and conversation, however this does not affect final choice.

7.6 Discussion For Experiment 2

Below is a summary of the results as they pertain to the predictions made by the individual theories.

1. Direct Access View

This view (Gibbs, 1986; Gibbs, 2002) did not receive empirical support from the experiment presented. Sarcastic Prosody interacted with Frequency in the Balanced homophone reaction time data, indicating that it was being processed with a different time frame than a Sincere interpretation. Due to the lack of support for this framework within the analysis, the remaining experiment focused on testing Graded Salience and Relevance Theory, which did each receive some support in the current experiment.

2. Graded Salience Hypothesis

Using this hypothesis, (Giora, 1997), several testable predictions were generated at the onset of the experiment. These predictions followed the notion that salience would be added to the interpretations in the case of Sarcasm interacting with previous factors, when discourse interpretations continually compete until participants are forced to make a choice. The previous cues that Sarcasm was considered to interact with included Frequency, Dominance, and Social Contextual information.

Biased Homophones: Graded Salience Hypothesis

When examining processing of Biased homophones using the eye-movement record, no effect of Sarcasm was seen within the data, either in the reported analyses or in additional behavioral analyses that were carried out including a direct test of Sarcasm as interacting with Frequency and Dominance. Instead, the effects of Frequency and Dominance alone persisted in directing attention and choice throughout the spoken discourse and decision of applicable context. This suggests that Sarcasm does not contribute salience given these conditions, and in order to be interpreted within the Graded Salience Framework, one would argue that Sarcasm is not viewed as a salient cue.

Balanced Homophones: Graded Salience Hypothesis

Within the eye-movement record, when tested Sarcasm did not interact with Frequency throughout as originally predicted (when tested using alternative models). Instead, during the processing of Sentence 2, it interacted with variables such as Change, generating a shorter fixation duration when there will be a shift in viewing contextual depictions, compared to sincere prosody (see Figure 7.6), as well as with Social

Contextual Ratings scores, increasing fixation time as the score increased. Together, this suggests participants are considering the Sarcasm they are presented with in the context of the overall conversation and its social implications, rather than as a highlight for specific lexical information (such as Frequency classification) as processing unfolds. However, this interaction does support the Graded Salience (Giora, 1997) prediction of being considered a salient cue given Balanced homophone conditions. However, given Sarcasm did not interact with Frequency, the interpretation is not without issue.

Behavioral: Graded Salience Hypothesis

Within the behavioral data, there was again no effect of Sarcasm in the Biased homophones (only Frequency and Dominance), thus it appears that the Bias or Dominance itself (Dominant or Subordinate meaning) interacts with Frequency and is deemed sufficient information to base choice of interpretation on. Or, another way to view these results in support of a Graded Salience approach (Giora, 1997), is that Dominance information contributes sufficient salience to choose an interpretation, and renders Sarcasm unnecessary. This could be an effect of the fact that Sarcasm does not directly interact with the mention of the homophone in the design.

In the case of Balanced homophones, a poverty of lexical information associated with the homophone, appeared to increase reliance on the Sarcasm as an informational cue soon after it was offered in Sentence 2. This was demonstrated through to final processing, as measured by reaction times. In reaction time measures, Sarcasm did interact with Frequency, (as was predicted to also occur within the eye-movement measures, according to Graded Salience). Additionally, the interaction does appear to support Graded Salience Framework, as it decreases reaction time when selecting a High

Frequency alternative, and increases reaction time when selecting a Low Frequency alternative (and removes the significant difference between the conditions) as compared to Sincere prosody, thus reversing the pattern. This suggests that Sarcastic Prosody did interact with the Frequency information on some level, even though it was not clear in the eye-movement record.

In summary, this data suggests that Sarcasm is considered useful in particular situations in the contexts created within the current experiment. As Sarcasm did not affect final choice for either Biased or Balanced homophones, it appears that as a salience cue, while in some cases it does appear to highlight an alternative, the alternative only receives additional consideration. Together, this results in partial support of the Graded Salience Framework as an explanation for Sarcasm resolution.

3. Relevance Theory

This framework (Sperber & Wilson, 1986; Sperber & Wilson, 1995) also predicted an interaction between factors Dominance, Frequency, and Sarcasm when choosing an interpretation of the homophone.

Biased Homophones: Relevance Theory

In the case of Biased homophones, when considering the lexical information associated with the homophone (Frequency, Dominance status), it appears that these cues create a discourse interpretation that is difficult to overcome. Indeed, within the data, it appears that listeners do not require sarcasm to resolve the ambiguous reference, even when the cue is available. However, Relevance Theory did predict that this was a likely outcome, since the information provided by Dominance and Frequency was strong. Here, Sarcasm was seen as irrelevant. Perhaps if the time course offered for resolution is

altered in the case of Biased homophones (as in Experiment 3), sarcasm can serve as a stronger contextual effect.

Balanced Homophones: Relevance Theory

Relevance Theory framework (Sperber & Wilson, 1986; Sperber & Wilson, 1995) predicts that only in some cases will sarcasm be beneficial. Here, it is only in the case of Balanced homophones where we see a strong effect of sarcasm as an informational cue being considered by listeners. The effect is observable from when Sarcasm is introduced in Sentence 2 (such that it interacts with Social Contextual Ratings, increasing processing of highly rated meanings), through to reaction time data (where it reverses the choice time pattern by Frequency seen in the Sincere prosody), as reported in the previous section. This singular effect of aiding in one context (Balanced homophones) was predicted from the Relevance Theory Framework.

Behavioral: Relevance Theory

The null effect of sarcasm in reaction times and final choice of Biased homophone interpretations is predicted by Relevance Theory by the prediction that the cue is not relevant enough to listeners. In the case of Balanced homophones, we see an interaction of Sarcasm and Frequency in reaction time data (again, such that time to choose between High and Low frequency alternatives is lower in the High Frequency condition and longer in the Low Frequency condition than the Sincere control). Here, the information is considered relevant, considering the lack of other available information. While Relevance Theory would predict differences between the final choice proportions in Balanced homophone conditions compared to Sincere Prosody, this was not observed.

Summary

Thus, as described above, it appears that both Graded Salience (Giora, 1997) and Relevance Theory (Sperber & Wilson, 1986; Sperber & Wilson, 1995) receive some support from the current experimental findings. In the case of Biased homophones, there was no observed effect of Sarcasm. While predicted by Relevance Theory, it is unclear whether sarcasm has no effect on highlighting alternatives for interpretation (given the fairly consistent Frequency by Bias interaction), or the design simply did not render sarcasm *important* enough to be considered, given listeners have access to other information and conversations were short. Thus, it is possible the design encouraged ambiguity to be resolved in Biased homophones based solely on Frequency and Dominance information because the homophone was not introduced with Sarcasm. Both Graded Salience and Relevance Theory predicted increased processing of one alternative in the Balanced condition, and it was here we observed the increase of processing meanings which were considered more Socially Contextually relevant.

When a homophone is directly mentioned using Sarcasm, the likelihood a speaker intends a different interpretation to be considered is elevated. Thus, the design of Experiment 3 introduced this manipulation, and also reduced the time from the introduction of the homophone to the decision on a context, to encourage listeners to utilize the prosody information further as a guiding heuristic in judgment of context. Additionally, this further differentiates Graded Salience and Relevance Theory approaches to sarcasm resolution.

CHAPTER 8

EXPERIMENT 3

8.1 INTRODUCTION

Experiment 3 aimed to further clarify the effect of Sarcasm on resolving ambiguous references in discourse by highlighting alternative interpretations. Particularly, the question of how Sarcasm interacts with other information given in a discourse, by continuing to investigate whether sarcasm can contribute to resolving ambiguity when it *highlights* the ambiguity by its position as part of the utterance. In contrast to Experiment 2, the current experiment utilized the same speaker to mention a homophone in an utterance that was spoken either Sarcastically, or Sincerely. Thus, instead of Sarcastic Prosody in Sentence 2 *following* the ambiguity generated by a speaker using a homophone in Sentence 1, here the second speaker was responsible for both the homophone and Sarcasm, leaving the listener (participant) to interpret intended meaning. This removes a presumption of Experiment 2, that participants may have held the expectation that the second speaker who utilized sarcasm, was aware of the first speaker's intended meaning of the homophone.

Previously introduced work by Blunter & Sommer (1998) identified that when a homophone is placed in focus using accent patterns, it is more likely to receive additional detailed semantic processing. Thus, it is expected that when the homophone is said sarcastically, the effects will differ from the Experiment 2 due to the increase in detailed lexical processing highlighting different information. If indeed it was the case that

sarcasm was not seen as a salient enough cue or a powerful enough context effect when Biased homophones were presented in Experiment 2, it is possible that in the current experiment, the use of Sarcasm in Sentence 2 to focus the homophone will change its perceived salience, when combined with the reduced amount of processing time, as listeners were asked for their interpretation directly after the sentence.

This difference also created a change in the Social Context being investigated. Experiment 3 focused on a situation where sarcasm is intended to serve a purpose for the listener directly. Within the experiment, the speaker using sarcasm has intentionally offered additional information in the form of Sarcasm to aid toward resolution of the discourse. For this reason as well, it is likely that the effects of given information observed in the current experiment will differ from those recorded in Experiment 2.

Focusing on the information provided by the homophones within the new design, when homophones involve Bias (a Dominant meaning), it is unclear if the resolution of the meaning will change from Experiment 2 when the word is embedded in the presentation of Sarcastic Prosody. The analysis of previous work (Blunter & Sommer, 1998) would lead to the hypothesis that since the prosody manipulation is now directly applied to the same utterance that contains the homophone, it should serve to heighten further the effects of lexical processing found in Experiment 2.

8.2 Methods For Experiment 3

Materials

In Experiment 3, the 30 remaining 2-sentence long conversations previously normed in Experiment 1 were utilized in conjunction with the same visual displays and filler items from Experiment 2. The discourses used the same homophones to introduce

ambiguity during the conversation. A sample version of the verbal stimuli specific to the current experiment can be viewed in Table 8.1, with all discourses listed in Appendix B. After each conversation, participants were asked to click on the picture depicting the context that they believed the conversation referred to.

Table 8.1. Experiment 3 Sample discourse.

Sentence	Speaker	
1	A	I feel like I have to buy these every year.
2	B (Sincere/Sarcastic)	Maybe you just don't know how to take care of a bulb properly.
?	(Text appearing after scene)	Which picture best fit the topic of conversation?

**B presented in Sarcastic Prosody in sarcastic conditions.*

8.3 Linking Hypothesis & Theoretical Predictions

Linking Hypothesis

Similar to the Linking Hypothesis for Experiment 2, the Linking Hypothesis for the current experiment also focuses on examining eye-movements at the offset of a key linguistic event within the spoken conversation stimuli as measures of processing of the events. This is a short time window immediately following the onset of the homophone within Sentence 2, where the additional processing caused by Sarcastic Prosody is considered to have the greatest effect on the processing of alternative meanings. Both the entire time window, and first fixations within the time window will be considered.

Additionally, as in Experiment 2, the impact of sarcasm can also be examined by looking at the overall sentence interpretation, in which choice proportions per meaning (one contextual representation versus the second), and reaction times to make a decision can be compared by factor condition.

Predictions

1. Direct Access View

Due to the lack of support from the behavioral reaction time models in Experiment 2, no additional hypotheses were made regarding the outcome of the current experiment that differ from those offered in Experiment 2[2] for the Direct Access View (Gibbs, 1986; Gibbs, 2002).

2. Graded Salience Hypothesis

Graded Salience predicts that cues are continuously being added to the calculation of salience of interpretations if they are deemed important (Giora, 1997). Thus, the predictions for Biased and Balanced homophones again follow from this general observation.

Biased Homophones

From this perspective (Giora, 1997), in the case of Biased homophones, the lexical information that is activated upon mention (Frequency and Dominance) should again contribute salience to the interpretation choice. If Sarcastic Prosody is used to mention the homophone, it should further increase the processing of the activated

---

[2] In addition, similar hypotheses are in effect for the current experiment in regards to the Muting Hypothesis, as well as the Echoic Reminder Theory, generally that the sarcasm may be interpreted as a reference to some distant social contexts, and therefore have an effect.

meanings. However, the alternative interpretation from Experiment 2's baseline (which was High Frequency, Dominant items) is more likely to be considered, given the direct application of prosody elevating the likelihood it will be considered a salience cue. Here, Sarcastic Prosody highlights the homophone, and the prosody is useful and informative in grading salience between the competing interpretations. Thus, the effect of Sarcasm is itself expected to have an impact on the salience of the competing interpretations, interacting with the Frequency and Dominance information after homophone onset. Additionally, participants have less time between hearing a direct mention of the homophone and indicating a meaning, Thus, eye-movement models of fixation patterns at homophone onset and shortly following should show the effect of Sarcasm quickly.

Balanced Homophones: Graded Salience

Once again, given Graded Salience, in the case of a Balanced homophone, just as in Experiment 2, Frequency effects are expected to interact with the processing of the Sarcastic utterance. However, at the late stage in the discourse, a greater salience boost should be given to the interpretation that is less likely, given it is the speaker using the Sarcastic Prosody. The Sarcastic Prosody should be viewed as a salience cue intended by the speaker to reverse a literal interpretation, or in this case, an initial interpretation.

Final Choice Patterns: Graded Salience

Graded Salience would predict the Sarcastic Prosody should create a clear effect of Sarcasm on final choice patterns, including both a main effect and interaction with covariates within reaction time data. Additionally, increased lexical processing of the homophone caused by sarcasm so close to the decision time, should result in slower reaction time data for the sarcasm condition.

3. Relevance Theory

Relevance Theory (Sperber & Wilson, 1995) would also make different predictions than Experiment 2 regarding the effect of sarcasm on processing, given later use of the homophone during the discourse. If the prosody is considered relevant at the late processing stage, the framework would predict additional lexical processing taking place because the homophone itself may now be considered "more important" or "marked" information in the discourse. Thus, it should serve to highlight information about the homophone that was previously not considered. As Relevance Theory would predict that once an interpretation is chosen, cues that do not agree with that interpretation are not necessarily deemed relevant if they do not serve as a contextual effect, we should again see an effect of sarcasm only in relation to certain contextual depictions of homophone meanings.

Biased Homophones: Relevance Theory

When considering contextual effects toward the end of the discourse where the homophone is introduced, Sarcasm should serve to highlight information that was previously not focused on. Sarcasm, to be considered relevant, would likely need to serve as a contradictory contextual effect, highlighting a different interpretation than was the literal interpretation considered before (it should indicate the opposite of the initial baseline in Experiment 2, which was Subordinate, High Frequency meanings). Thus, the Dominant meanings of Biased homophones should receive additional processing, or those that are Low in Frequency.

Balanced Homophones: Relevance Theory

If at the late stage of the discourse, Sarcasm combines with the mention to produce increased lexical processing, it may serve to highlight information that was previously not as salient, such as the Frequency factor in the case of Balanced homophones. As Frequency and Social Context are the only cues provided, there should be increased activation of this information, given the focused mention, thus an interaction of it with Sarcasm.

Final Choice: Relevance Theory

As in Experiment 2, Relevance Theory would only predict an increase in final choice reaction time data if Sarcasm is considered relevant and requires processing, and this increase should also then be accompanied by changes in final choice. If there is no change in final choice pattern, then while increasing processing of alternatives, sarcasm has had a detrimental effect by causing additional processing with no payoff.

Summary of Predictions

Within the current experiment, it is expected that effects will be greater at homophone offset than in Experiment 2, due to increased lexical processing caused by the Sarcastic Prosody. The hypotheses presented above pertain to Sarcasm resolution in the context presented in Experiment 3, and rely on the assumption that listeners are able to integrate multiple information streams quickly and prioritize to resolve the homophone ambiguity in the discourse. A goal of the current experiment is therefore to determine whether Sarcasm can affect the processing of Biased homophones to a greater extent, by highlighting an alternative to the baseline interpretation in Experiment 2, while still affecting the processing of Balanced homophones.

8.4 Results For Experiment 3

Reporting of the results for the current experiment will proceed in the same format as the previous experiment. Data preparation was carried out in the same manner as well. As a note, in response to the outcome of the Experiment 1C regarding the auditory items, models were run without an item (item 4). This left a total of 29 experimental items.

8.4.1 Eye-movement Data

The analysis for Experiment 3 focused exclusively on Sentence 2, as there were no linguistic events of interest for the current research question in Sentence 1. The goal of Sentence 1 was to engage a listener in the conversation before the Sarcastic Prosody and mention of the homophone. The analyses are once again split by Bias (Dominance) of meanings with analysis of Biased, followed by Balanced homophones.

Sentence 2 Biased Homophones

For the Biased homophones, at homophone onset +750 *ms*, the model analyzed corresponds to [11]. There was a moderately significant interaction of Sarcasm with Frequency ($p = 0.08$), such that in the Sarcastic condition, Frequency had no effect on the length of fixations, while Low Frequency items received shorter fixations in the Sincere Prosody condition, (see Figure 8.1). Additionally, there was a main effect of the Picture Proportion (moderately significant, $p = 0.06$) and Frequency ($p < 0.03$, High Frequency items were fixated on longer), indicating that participants were focusing attention after hearing the homophone according to readily available visual, lexical, and prosody information. However, Dominance was not a significant factor, nor did it interact with Frequency when tested in a separate model (see Table 8.2 for model information). This

may indicate that sufficient context had been presented within the sentence to remove the initial biasing effect of Dominance, or that participants had guessed the homophone by this point, and the mention just increased processing of some characteristics. Faux Pas task responses were tested in separate models for Experiment 3, but as they did not significantly alter the results (and did not contribute as significant predictors) they were removed from the analysis for clarity. For a clear comparison to the Experiment 2 Sarcasm data in the first portion of Sentence 2 (where the homophone had already been heard), Figure 8.2 divides fixation durations by Prosody, Frequency, and Dominance; however this was not a significant interaction in the current experiment.

[11] $Y_{ij} = \beta_{0i} + \beta_1*\text{PictureProportion}_{1ij} + \beta_2*\text{SocialContextRating}_{2ij} + \beta_3*\text{Sarcasm}_{3ij} + \beta_4\text{Frequency}_{4ij} + \beta_5*\text{Dominance}_{5ij} + \beta_6*\text{SocialContextRating}*\text{Sarcasm}_{6ij} + \beta_7*\text{SocialContextRating}*\text{Frequency}_{7ij} + \beta_8*\text{Sarcasm}*\text{Frequency}_{8ij} + \beta_9*\text{SocialContextRating}*\text{Sarcasm}*\text{Frequency}_{9ij} + b_{i1}*\text{Subject}_{1i} + b_{i2}*\text{Item}_{2j} + \varepsilon_{ij}$

The results of the current analysis are of interest because they indicate that when Sarcastic Prosody is used with a homophone referent, there is a different level of activation of alternatives produced by the lexical access than when Sarcastic Prosody is not used, as in Experiment 2. Additionally, within the Biased homophone analyses of Experiment 2, the Sarcastic Prosody factor was never significant within the eye-movement data, nor did it interact with any other effect. The significance of the factor interacting with Frequency in the current analysis suggests that the different context created by the conversation did successfully alter the listener's expectations of the discourse.

A second analysis focused on the first fixation duration beginning after homophone onset found a moderately significant effect of Social Rating ($p = 0.08$), such

that items with higher Social Contextual Ratings received *shorter* fixations, as well as a main effect of Frequency ($p < 0.007$), such that Low Frequency items also received shorter fixations. Additionally, there was also an interaction of Social Contextual Rating and Frequency, and Sarcastic Prosody and Frequency, such that given Sarcastic Prosody, the effect of Frequency was null (p's < .05) (see Table 8.3 for model information). The Sarcasm by Frequency interaction can be viewed in Figure 8.3. Finally, when Dominance was tested in a different model, again it did not interact with either Social Context Ratings, Sarcastic Prosody, or Frequency, so it was tested only as a main effect within the model as indicated in [11].

Table 8.2. Models of fixations to Biased homophones in Sentence 2, full 750 *ms*.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, Sarcasm, Frequency, and Dominance. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes Sincere, High Frequency, Dominant reference group. Significant estimates are bolded.

Biased Homophones full 750 ms time window

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	269.01	12.17	22.10	**0.001**
Picture Proportions	160.69	86.31	1.86	**0.06**
Social Contextual Rating	-19.35	16.31	-1.19	= 0.24
Sarcastic Prosody	0.96	14.15	0.07	= 0.95
Low Frequency	-29.73	12.80	-2.32	**0.03**
Subordinate Meaning	-0.09	11.05	-0.01	= 0.99
Social Rating*Sarcastic	19.53	21.94	0.89	= 0.37
Social Rating*Low Frequency	32.95	19.39	1.70	**0.09**
Sarcastic*Low Frequency	31.46	17.97	1.75	**0.08**
Social Rating*Sarcastic*Low Freq.	-16.83	28.57	-0.59	= 0.56

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	501.47	22.39
Item	(Intercept)	86.19	9.28
Residual		9185.84	95.84

Model Fit Measures	
AIC	6585
BIC	6641
logLik	-3280

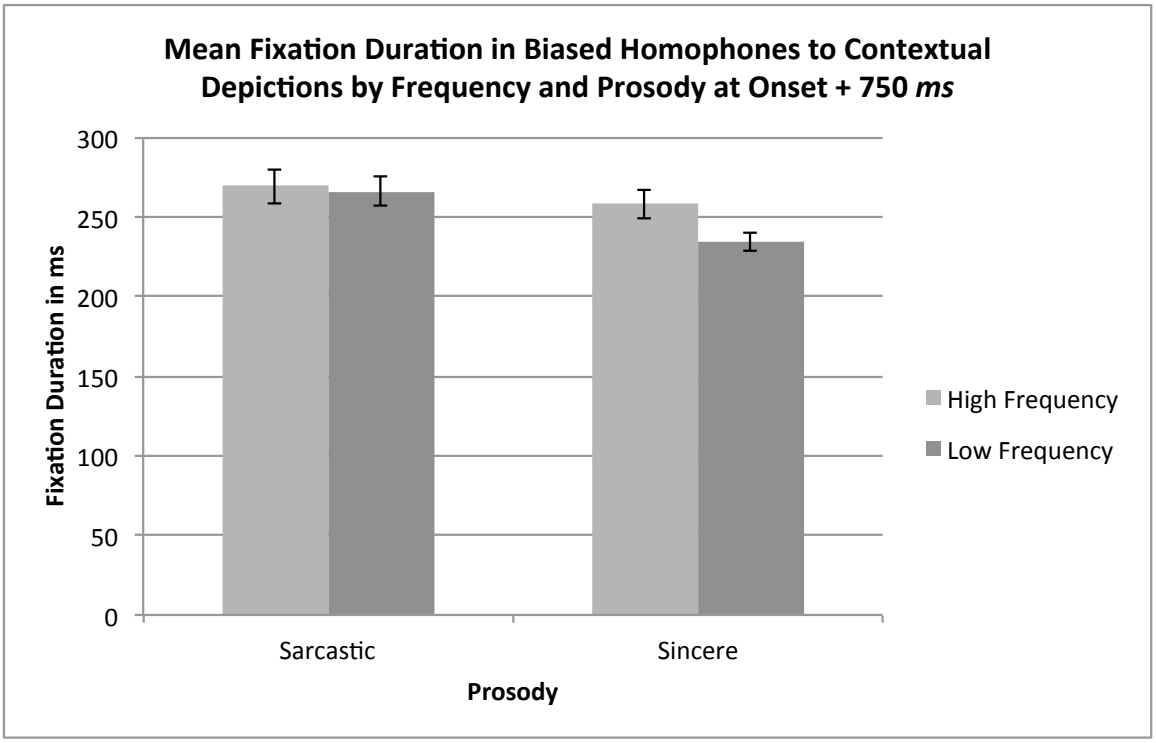

Figure 8.1. Mean Sentence 2 fixation durations in *ms* at homophone onset to Biased homophones by Sarcasm and Frequency. Standard error of mean depicted.

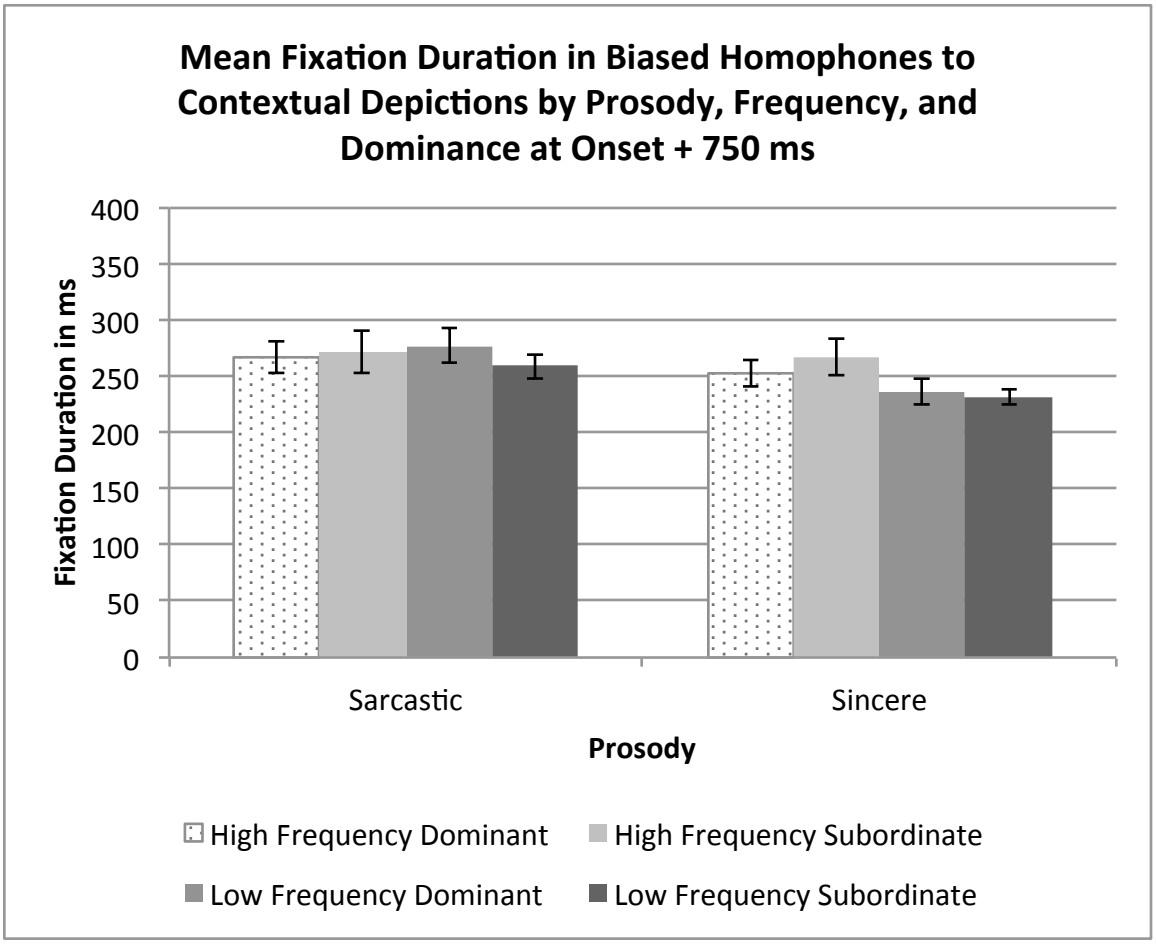

Figure 8.2. Mean Sentence 2 fixation durations in *ms* at homophone onset to Biased homophones by Sarcasm, Frequency, and Dominance. Standard error of mean depicted.

Table 8.3. Models of first fixations to Biased homophones in Sentence 2.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, Sarcasm, Frequency, and Dominance. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes Sincere, High Frequency, Dominant reference group. Significant estimates are bolded.

Biased Homophones first fixation duration only after homophone onset time

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	293.34	47.67	20.00	**0.001**
Picture Proportions	140.25	105.50	1.33	= 0.18
Social Contextual Rating	-36.14	20.30	-1.78	**0.08**
Sarcastic Prosody	-10.90	18.23	-0.60	= 0.55
Low Frequency	-43.01	15.83	-2.72	**0.007**
Subordinate Meaning	-11.39	13.99	-0.81	= 0.42
Social Rating*Sarcastic	31.33	28.40	1.10	= 0.27
Social Rating*Low Frequency	53.04	24.74	2.14	**0.04**
Sarcastic*Low Frequency	44.70	23.26	1.92	**0.06**
Social Rating*Sarcastic*Low Freq.	-27.62	36.39	-0.76	= 0.45

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	323.81	18.00
Item	(Intercept)	0.00	0.00
Residual		10915.97	104.48

Model Fit Measures	
AIC	4501
BIC	4552
logLik	-2237

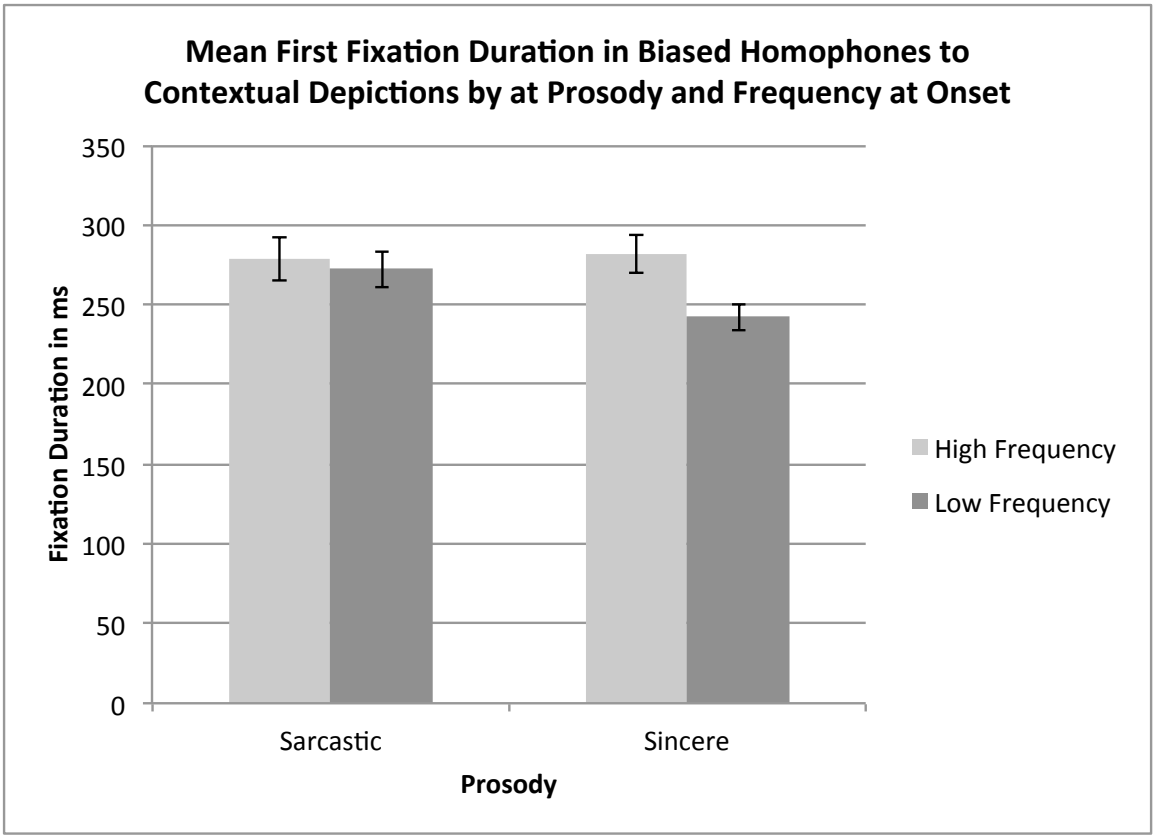

Figure 8.3. Mean first fixation duration to Biased homophones by Prosody and Frequency at homophone onset in Sentence 2. Standard error of mean depicted.

Sentence 2 Balanced Homophones

Turning to the analysis of the Balanced homophones in the same 750 *ms* time window, [11] was modified and the factor Dominance was removed from the equation. However, other than this change, the same two analyses were performed. Table 8.4 reports the full results of this analysis. Here, none of the predictors proved significant. This suggests that at homophone onset and immediately after, listeners were not focusing on any one informational cue to direct attention in the case of a Balanced homophone.

When the model was restricted to first fixation duration data only, the estimates were also insignificant (see Table 8.5).

Table 8.4. Models of fixations to Balanced homophones in Sentence 2, full 750 *ms*.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, Sarcasm, and Frequency. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes Sincere, High Frequency reference group. Significant estimates are bolded.

Balanced Homophones full 750 ms time window

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	247.92	19.16	12.94	**0.001**
Picture Proportions	47.59	116.83	0.41	= 0.68
Social Contextual Rating	7.12	27.88	0.26	= 0.80
Sarcastic Prosody	17.15	30.39	0.56	= 0.57
Low Frequency	11.86	20.75	0.57	= 0.57
Social Rating*Sarcastic	7.49	42.98	0.17	= 0.86
Social Rating*Low Frequency	-2.27	35.58	-0.06	= 0.95
Sarcastic*Low Frequency	-25.67	34.31	-0.75	= 0.45
Social Rating*Sarcastic*Low Freq.	-4.07	54.67	-0.07	= 0.94

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	0.001	0.35
Item	(Intercept)	0.00	0.00
Residual		10458	102.26

Model Fit Measures	
AIC	4042
BIC	4088
logLik	-2009

Table 8.5. Models of first fixations to Balanced homophones in Sentence 2.

Data modeled according to predictors Experiment 1 Picture Proportion data, Social Contextual Rating, Sarcasm, and Frequency. Models are of fixation durations, with subjects and items containing random intercepts. Intercept estimate includes Sincere, High Frequency reference group. Significant estimates are bolded.

Balanced Homophones first fixation duration only after homophone onset time

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	267.25	25.89	10.33	**0.001**
Picture Proportions	78.89	141.95	0.56	= 0.58
Social Contextual Rating	-4.96	38.70	-0.13	= 0.90
Sarcastic Prosody	2.26	37.96	0.06	= 0.95
Low Frequency	0.22	28.24	0.01	= 0.99
Social Rating*Sarcastic	46.76	55.33	0.85	= 0.40
Social Rating*Low Frequency	3.18	47.23	0.07	= 0.95
Sarcastic*Low Frequency	-0.90	43.42	-0.02	= 0.98
Social Rating*Sarcastic*Low Freq.	-44.85	69.85	-0.64	= 0.52

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	1312.8	36.23
Item	(Intercept)	0.00	0.00
Residual		12306.0	110.93

Model Fit Measures	
AIC	2738
BIC	2779
logLik	-1357

Summary of Sentence 2 Eye-Movement Data

In summary, while hearing a homophone in Sarcastic Prosody contributed to processing Biased homophones, for Balanced homophones, effects were *reduced* as compared to Experiment 2. Thus, the current experiment with the modified conversations provided an apparent reversal of effects. The time window of analysis was not extended further in the current experiment because the homophone was spoken toward the end of the sentence. In order to determine whether sarcasm contributes to

processing of final choice, reaction time data and choice proportions were calculated as in Experiment 2 for the current experiment.

Again as a note, models were tested that included Sarcasm interacting with Dominance for the Biased homophones, but the models did not offer significant fit improvement nor were the effects significant, so the models reported here are those that are more tractable. Additionally, in light of the irrelevance of the Faux Pas data within the models in Experiment 2, although it was tested as a predictor in a more complex version of the results than is modeled here, it was insignificant in contributing to model fit and thus not reported in order to maintain tractability. The analysis of the behavioral data is presented below.

8.4.2 Behavioral Data

Reaction Times

Reaction times greater than 4000 *ms* were immediately removed from the analysis as they were considered outliers (6 responses, < 1%).

Biased Homophones

The same multilevel mixed-effects model with random intercepts for subjects and items with factors of Sarcasm (prosody) and covariates Frequency and Dominance (within the Biased homophones) was used to model reaction times (see [9] and [10] for models). In contrast to the results of the reaction time data in Experiment 2, within the Biased homophones, only Social Contextual Ratings significantly predicted reaction time, such that the higher the rating, the shorter the reaction time (see Table 8.6, Model 1). When Social Contextual Ratings were removed (see Table 8.6, Model 2), there were no significant predictors of reaction time, and it did not improve fit of Sarcastic Prosody.

Balanced Homophones

Within the Balanced Homophones, neither a model of reaction time containing Social Contextual Ratings, or without them, significantly predicted reaction times. However, the increase in the variability of response times when the sarcasm manipulation occurs late in the discourse, and is combined with the homophone reference, supports a likelihood of additional lexical processing due to the prosody (Blunter & Sommer, 1998). The fact that it occurred so late in the discourse suggests that lexical processing of the meaning is likely not yet complete, nor are there sufficient alternative cues when participants are forced to make a decision on which context fits the discourse, resulting in higher variability.

Table 8.6. Models of Experiment 3 reaction time data by Bias.

Data modeled by Bias according to predictors 1) Sarcastic Prosody and Social Contextual Rating 2) Sarcastic Prosody, with covariates Frequency and Dominance of homophone where applicable. Models are of reaction times, with subjects and items containing random intercepts. For Biased conditions, intercept estimate includes High Frequency, Dominant, Sincere condition as reference groups, while Balanced intercepts do not contain Bias. Significant effects are bolded.

Model 1. *Biased Homophones Sarcastic Prosody & Social Contextual Ratings*

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	995.81	101.47	9.42	**0.001**
Social Contextual Rating	-148.32	65.55	-2.26	**0.03**
Low Frequency	34.38	119.94	0.29	= 0.77
Subordinate Meaning	-97.24	127.54	-0.76	= 0.45
Sarcastic Prosody	5.46	97.31	0.06	= 0.96
Low Frequency*Subordinate	53.23	158.04	0.34	= 0.74
Low Frequency*Sarcastic Prosody	-27.25	133.20	-0.20	= 0.84
Subordinate*Sarcastic Prosody	212.52	152.16	1.40	= 0.16
Low Freq.*Subordinate*Sarcastic	-161.51	200.25	-0.81	= 0.42

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	56,464	237.62
Item	(Intercept)	22,536	159.80
Residual		325,904	570.88
Model Fit Measures			
AIC	9244		
BIC	9297		
logLik	-4610		

Model 2. *Biased Homophones Sarcastic Prosody Only*

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	881.83	93.46	9.44	**0.001**
Low Frequency	92.32	113.54	0.81	= 0.42
Subordinate Meaning	30.84	114.83	0.27	= 0.79
Sarcastic Prosody	4.21	97.83	0.04	= 0.97
Low Frequency*Subordinate	-27.98	153.13	-0.18	= 0.86
Low Frequency*Sarcastic	-13.90	133.78	-0.10	= 0.92
Subordinate*Sarcastic	194.38	152.79	1.27	= 0.20
Low Freq.*Subordinate*Sarcastic	-136.92	201.01	-0.68	= 0.50

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	57,867	240.56
Item	(Intercept)	18,122	134.62
Residual		330,151	574.59
Model Fit Measures			
AIC	9257		
BIC	9305		
logLik	-4618		

Model 3. *Balanced Homophones Sarcastic Prosody & Social Contextual Ratings*

Fixed Effects Coefficient Estimates	β *Est.*	*Std. Error*	*t*	*p* <
Intercept	883.29	131.99	6.69	**0.001**
Social Contextual Rating	-4.33	11.15	-0.04	= 0.97
Low Frequency	85.76	134.35	0.64	= 0.52
Sarcastic Prosody	173.26	126.60	1.37	= 0.17
Low Frequency*Sarcasm	-149.82	154.84	-0.97	= 0.33

Random Effects	*Name*	*Variance*	*Std. Dev.*
Subject	(Intercept)	1312.8	36.23
Item	(Intercept)	0.00	0.00
Residual		12,306.0	110.93
Model Fit Measures			
AIC	2738		
BIC	2779		
logLik	-1357		

Model 4. *Balanced Homophones Sarcastic Prosody Only*

Fixed Effects Coefficient Estimates		β *Est.*	*Std. Error*	*t*	*p* <
Intercept		882.82	125.92	7.01	**0.001**
Low Frequency		86.23	128.31	0.67	= 0.50
Sarcastic Prosody		172.45	125.57	1.37	= 0.17
Low Frequency*Sarcasm		-149.85	154.21	-0.97	= 0.33
Random Effects	*Name*	*Variance*	*Std. Dev.*		
Subject	(Intercept)	81,145	284.86		
Item	(Intercept)	31,864	178.51		
Residual		354,874	595.71		

Model Fit Measures	
AIC	4668
BIC	4694
logLik	-2327

Final Choice Data

Finally, proportions of chosen responses by condition were analyzed (answers were quantified into the six conditions previously noted for the eye-movement data, e.g., "Balanced – Low Frequency", etc.) and a Chi-Square test for Independence was completed on the resulting proportions. The test found no differences by condition Sarcasm. However, when comparing the differences between conditions to those found in Experiment 2, the data does *trend* toward having differences between conditions, particularly when we focus on the Low Frequency conditions (see Table 7.7 for comparison). Additionally, there appears to be the same overall effect of Frequency in choice selection, a replication of Experiment 2. To view the proportions, see Table 8.7.

Table 8.7. Experiment 3 final choice proportions by factor Sarcasm.

Contextual Depiction Chosen	Sincere	Sarcastic
Balanced High Frequency	11.33%	11.19%
Balanced Low Frequency	24.49%	21.91%
Biased Subordinate Low Frequency	20.41%	15.15%
Biased Subordinate High Frequency	12.24%	11.66%
Biased Dominant Low Frequency	15.42%	21.91%
Biased Dominant High Frequency	16.10%	18.18%

8.5 Discussion For Experiment 3

The differences from Experiment 2 observed in the current experiment suggest that Sarcasm is able to serve multiple purposes when being used by a speaker, with the intention of highlighting alternative interpretations, or serving a beneficial role. The conditions tested in Experiment 3 also added additional information to the comparison of the theoretical perspectives, and their ability to account for sarcasm processing differences in discourse resolution. Combining the sarcasm and homophone into one sentence allowed the Sarcastic Prosody to highlight the ambiguity further, to determine the effect on the lexical access process. The differences presented in the models from Experiment 3 versus those from Experiment 2 suggest that as expected, there are differences in the utilization of Sarcastic Prosody in processing ambiguous references according to the constraints of the discourse and task (placement and timing). These differences will be discussed further within the General Discussion. Returning to the discussion of the theoretical perspectives as they pertain to the current experiment, the specific hypotheses are addressed.

1. Direct Access View

Briefly, since there were differences identified in Experiment 2 suggesting that Direct Access View did not receive support from that experiment, the only finding noted here is the response time data in the current experiment *did* offer some support of Direct Access View, given the constraints presented for the task. This support was based on the fact that there were not significant differences in response times predicted by Sarcasm in the Biased or Balanced homophone conditions, however Social Context did affect reaction time when included in the model for Biased homophones (Gibbs, 1986; Gibbs, 2002)[3].

2. Graded Salience Hypothesis

Biased Homophones: Graded Salience Hypothesis

Within the Graded Salience framework (Giora, 1995; Giora, 1997), in the eye-movement data, in the case of Biased homophone usage, the approach received partial support. There was a main effect of Frequency, however, Low Frequency alternatives were considered less overall (see Table 8.2). Additionally, there was a Sarcasm by Frequency interaction (see Figure 8.1) such that in the Sarcastic Prosody condition, the Low Frequency alternative received additional processing time than in the Sincere condition, and processing time that was not significantly different from the processing of High Frequency alternatives in the Sarcastic condition. However, Dominance did not interact with Sarcasm (see Figure 8.2). Thus, it appears Sarcastic Prosody added salience as measured by increased processing to Low Frequency interpretations,

---

[3] Although not directly tested, Echoic Reminder Theory and Muting Hypothesis (Kreuz & Glucksberg, 1989, Dews & Winner, 1995) did not receive readily applicable support.

regardless of Dominance. This suggests that the conversational manipulation was successful, as there was a strong effect of Dominance in Experiment 2 at homophone onset when Sarcastic Prosody was not used to present the homophone.

Balanced Homophones: Graded Salience Hypothesis

Alternatively, there was no effect of Sarcasm on the processing of the Balanced homophones post onset (either in first fixation data or through a longer time window), suggesting that Sarcasm was not considered a strong enough salience cue at this point. It's possible that listeners were activating (or attempting to) activate other information to guide interpretation.

3. Relevance Theory

Biased Homophones: Relevance Theory

Once again, Relevance Theory (Sperber & Wilson, 1986; Sperber & Wilson, 1995) also receives support. Within the Biased homophone eye-movement models, Sarcastic Prosody interacted with Frequency such that there was additional processing given Sarcastic Prosody to the Low Frequency alternative, compared to Sincere Prosody. However, the interaction was also such that there was no difference between the processing of High and Low Frequency alternatives, suggesting it served as a contradictory contextual effect, increasing fixations in the Low Frequency conditions, as the High Frequency alternatives were being considered to a greater extent.

Balanced Homophones: Relevance Theory

In terms of Balanced homophones, the lack of results in both the eye-movement and behavioral data can be reduced to Sarcasm not being viewed as a relevant cue given the conversational constraints. Instead, listeners may have focused on the lexical level

information highlighted by the prosody, perhaps searching for Dominance information that was unavailable.

Behavioral: Graded Salience Hypothesis & Relevance Theory

Sarcasm had no effect on the reaction time data. Within the final choice data, the effect of Sarcasm did not persist, however the final choice proportion data do differ from Experiment 2. There was expected to be a difference in both reaction time and final choice according to Graded Salience and Relevance Theory, so the lack of the effect does not provide any support for either theory.

Summary

In the current experiment, within the eye-movement record the effects of Sarcasm observed were exclusively in Sentence 2, directly after homophone onset for the Biased homophones. These effects were in the form of an interaction with Frequency information, such that Low Frequency information received additional processing in the Sarcastic, compared to Sincere Prosody conditions. Effects of Social Contextual Rating also interacted with Frequency and Sarcasm in the Biased homophone condition. For Low Frequency items, higher Social Contextual Ratings increased processing time. Additionally, there were no effects of Sarcasm on the processing of Balanced homophones within the current experiment in any measure.

These findings, along with the lack of differences in the prediction of reaction time data, suggest that the conversational constraints that were manipulated across experiments effectively changed how Sarcasm was utilized by listeners to activate alternative interpretations as they resolved ambiguous homophone references. In the current experiment, the presence of Sarcasm to highlight the homophone itself appeared

to increase lexical processing of some Experiment 2 Sentence 1 baseline information for Biased homophones, namely Frequency classification. Alternatively, for Balanced homophones, no covariate nor predictor immediately contributed to the processing of alternative meanings (see Table 7.5, Model 2). The implications of these findings are examined further in the General Discussion that follows.

CHAPTER 9

GENERAL CONCLUSIONS

The current study aimed to demonstrate that sarcasm could be considered useful information when choosing between multiple discourse interpretations, and that speakers can potentially *utilize* sarcasm to highlight a particular alternative. To do so, listeners must be able to demonstrate preference for selective interpretations, when sarcasm is used. In order to examine this question in more detail, homophones were used to introduce ambiguity into a series of spoken conversations. These conversations were generated by altering when the homophone ambiguity was introduced, and whether the Sarcastic Prosody was in the same utterance, thus altering a listener's expectations about the speaker's intentions within the conversations they were hearing. This work also served to begin a wider examination of the specific purposes that sarcasm can serve, while situating this examination using compatible existing theories of sarcasm processing.

The current study was able to manipulate the effectiveness of sarcasm in discourse, such that given different Situational Contexts, Sarcastic Prosody was more or less effective at highlighting alternative interpretations of homophones. The process of resolving sarcasm has been addressed in multiple theories currently found within the literature. These theories of sarcasm processing have traditionally been applied to statements situated within a variety of contexts in conversation and discourse. A subset of these theories have been applied to specific topics such as the placement of a

compliment or censure in conversation (Dews & Winner, 1995) or the use of sarcasm to reference social norms in conversation (Pexman & Olineck, 2002, Kreuz & Glucksberg, 1989; Gibbs, 2002). However, a few are also capable of describing language processing more generally (Giora, 1997; Sperber & Wilson, 1995). Both Graded Salience (Giora, 1997) and Relevance Theory (Sperber & Wilson, 1995) fall into the latter category, and were tested specifically within the framework of the current study, which utilized a VWP design and focused on examining claims generated from these frameworks using Linking Hypotheses. The implications of the eye-movement studies are discussed first.

As a brief summary, Experiment 1 provided information on the covariates such as Social Contextual Ratings and visual display information, and provided a verification of the items for use in the VWP studies. The measurement of Social Contextual Ratings accounted for the fact that some of the homophones (e.g., "jeans") allowed easier use of a Sarcastic comment to be generated, as reported by the subject pool. Additionally, the visual display information was normed such that the contributions of particularly vivid pictures, if found, could be controlled for. Experiment 2 examined the effect of Sarcasm following the introduction of an ambiguous homophone by a speaker. In the case of Biased homophones, it was found that the Sarcastic Prosody cue was overlooked in favor of prior Dominance (which meaning was more likely to be activated given no context information) and Frequency (written corpus classification) information that interacted together. The Graded Salience Hypothesis did particularly well predicting the activation of Frequency and Dominance information, which it would identify as salience cues, indicating slightly stronger support for this approach than Relevance Theory (Giora, 2002). However, in the case of Balanced homophones, Sarcastic Prosody increased

processing of meanings based on Social Contextual information. Additionally, Sarcasm interacted with the Change variable, such that if listeners switched to fixate on an alternative meaning, the current fixation was shorter. Experiment 3 focused on the effect of Sarcasm when it introduced an ambiguous homophone following a neutral sentence. In this case, when a Biased homophone was presented, Sarcasm interacted with Frequency, such that Low Frequency alternatives received additional processing in the Sarcasm condition, as compared to the sincere condition. However, when Balanced homophones were used, none of the predictors effectively contributed to resolving the homophone reference.

Taking these results into account, first, the use of sarcasm during a discourse has an influence on the process of ambiguity resolution as measured by sustained attention to alternatives, compared by analyzing fixation durations. Differences in processing patterns (as measured fixation durations and first fixations) were observed given the presence or absence of Sarcasm following an ambiguous referent or highlighting it in an utterance. Additionally, depending on the combination of the Sarcasm and a direct mention of the homophone, Sarcasm processing served to highlight different attributes of the homophone (Frequency, Dominance, or the Social Context). Both the Graded Salience Hypothesis and Relevance Theory could reliably explain a portion of the findings utilizing the mechanisms provided by the theory, however neither completely accounted for the data (Giora, 1997; Sperber & Wilson, 1995). The data that was unaccounted for by both theories tended to be the Behavioral data. Neither theory predicted that across all conditions Sarcasm would fail to change the final choice. Additionally, Relevance Theory had trouble accounting for the lack of an effect of

Sarcasm in both the Biased and Balanced conditions of Experiment 3, as did Graded Salience Hypothesis.

The differences in the eye-movement patterns modeled within Experiments 2 and 3 offers a contribution to what is known regarding sarcasm resolution and report findings not previously found the literature. Specifically, as predicted initially, sarcasm *does* appear to reliably alter the consideration of potential discourse interpretations when present. For the homophones presented in the current work, in Experiment 2, when Sarcasm followed the mention of a Biased homophone, there was no observable effect of sarcasm on processing. Participants began by fixating in Sentence 1 according to the Frequency and Dominance information associated with a homophone when it was mentioned, and reverted to using this information in the second 2000 *ms* of Sentence 2. However, in the case of Balanced homophones, while in Sentence 1 there was no baseline effect of Frequency, the Low Frequency meanings did not receive additional processing in Sentence 2 when Sarcasm was used. Instead, Social Contextual information was activated. Thus, in Experiment 2, Sarcasm appeared to be used by participants as a piece of information that highlighted a reliance on the conversation itself and plausible contexts, given the poverty of clearly Dominant meanings, or strong Frequency information. Relevance Theory is able to account for this data pattern by arguing Sarcasm is only relevant in the latter context (Balanced Homophones) (Sperber & Wilson, 1995). The Graded Salience Hypothesis had more difficulty accounting for all differences in processing patterns.

In Experiment 3, where the mention of the homophone appeared late in the conversation (in Sentence 2) and was highlighted with Sarcastic Prosody, we observe a

reversal in the effect of sarcasm on processing preference. Again, here there was a strengthening effect of Sarcasm in the Biased homophone condition (with Sarcasm highlighting Frequency information), and no effect of Sarcasm in the Balanced homophone condition. Thus, for Biased homophones Sarcasm appears to serve as a salient, relevant factor when considering multiple competing interpretations in the given Social Context. Alternatively, in the Balanced homophones, during processing the poverty of information available appears to lead listeners to disregard the Sarcasm, and continue trying to resolve the intended reference without clear strategy. Once again, Relevance Theory (Sperber & Wilson, 1995) can account for these findings by arguing that Sarcasm is sufficiently relevant in the case of Biased homophones given decreased time from mention to a forced decision of interpretation. However, Graded Salience can account for the findings by arguing that Sarcasm becomes a salient cue in the case of interpreting Biased homophones, but is not sufficiently strong enough to differentiate itself given the other cues it competes with in the case of Balanced homophones.

When considering the behavioral data collected within the current work more closely, both reaction time data and the choice of a final contextual depiction by participants, the overall effect of Sarcasm varied. In the reaction time data, for Experiment 2, within the Biased homophones Frequency and Dominance information appeared to affect processing of reaction times, while in the Balanced homophones, Social Contextual Ratings, Frequency, and the interaction of Sarcastic Prosody with Frequency affected reaction times. These effects were such that given Sarcastic Prosody, High Frequency alternatives were responded to faster than Low Frequency alternatives, the opposite pattern that was seen in the Sincere Prosody conditions. However, in

Experiment 3, where there was significantly less time between the lexical introduction of the homophone and a forced choice of interpretation, only within the Balanced homophones did any predictor achieve significance, the Social Contextual Rating variable.

Within Experiment 2, the situational context was intended to signal to participants that the second speaker was aware of the intended homophone's meaning, while in Experiment 3, participants were intended to assume that the interlocutor of the second speaker may not be aware of intended meaning, but that the speaker had chosen sarcasm for a reason. The differences in the findings of the contribution of Sarcasm and Social Context to the reaction time findings gives credence to the theories introduced in the introduction of this work, which to varying degrees suggested that there is a Social Context component to Sarcasm processing. The null effect of the Social Rating information in the Biased homophones condition of Experiment 2, and the Balanced condition of Experiment 3 may be due to the different processing assumptions of the participant, leading them to consider alternative interpretations based on other forms of information. The observation that the selection of final choice is not significantly altered by Sarcasm suggests that regardless if it introduces additional salience, or is considered relevant in highlighting alternative explanations, within this paradigm its overall effect is weaker than the lexical level information modeled in the current work.

Summary

Thus, when we consider these findings in their entirety, what emerges is the view that sarcasm is used in discourse resolution by comprehenders as an additional marker of information importance. It appears that when given time to process sarcasm, participants

either consider sarcasm irrelevant when other strong cues (such as Dominance and Frequency) are available (see Tables 7.4 & 7.6), or it strengthens one interpretation (such as the interaction of Sarcasm with Social Context and Frequency, in the eye-movement and behavioral data of Experiment 2, see Table 7.5 & 7.6). However, when Sarcasm is used in the same statement as an ambiguous referent, if participants are given little time to process they are more likely to attempt increased lexical processing of the referent, and process characteristics that become activated at baseline to a greater extent (see Tables 8.2 – 8.5). Since the behavioral choice data did not reflect a change it appears that Sarcasm simply activated further an additional discourse interpretation, compared to Sincere Prosody. Future studies should consider this information on the contribution of Sarcastic Prosody to activating multiple interpretations, when utilizing conversations that contain utterances with Sarcastic Prosody.

In terms of what this suggests for the theoretical perspectives considered, we can conclude that there exists within both Graded Salience (Giora, 1997) and Relevance Theory Frameworks (Sperber & Wilson, 1995) sufficient mechanisms to begin explaining sarcasm processing, as it pertains to reference resolution. While neither framework perfectly accounts for the data presented, Relevance Theory does contribute to our understanding of sarcasm as a chosen linguistic mechanism, and aid in explaining why some individuals may have more difficulty processing sarcasm. It takes particular sets of contextual constraints for individuals to begin to consider Sarcasm as a relevant or salient informational mechanism, and it is likely that Sarcasm is not always sufficiently grounded within a conversation or discourse. Certainly, the current findings contribute to demonstrating how just how dependent the consideration of sarcasm during processing is

on contextual variables; such as the Social Context an utterance is placed within and the placement of the Sarcasm in the utterance itself. Thus, while the Faux Pas task as a measure of individual differences did not significantly contribute to the models within the current study, individual differences may be traced to some comprehenders of spoken language utilizing cues that were not present in this study to a differing extent than their peers, or requiring less "ground" for sarcasm. These variables could be examined in future work, and would likely further relate to the Social Contextual cues noted within the current work.

In conclusion, the current work first demonstrated and then clarified that sarcasm, when used by a speaker, can highlight a particular interpretation of a previously ambiguous homophonic statement, given proper placement within a conversation. In doing so, this study demonstrated that hearing sarcasm selectively generates additional consideration of multiple discourse interpretations by listeners. In addition, this project served to disentangle theoretical explanations of sarcasm, suggesting that those with broad approaches to language processing may be needed to account for its multiple effects within language. Additionally, the current work found that existing approaches may already have mechanisms available for explaining the processing of sarcasm in multiple conversational contexts. In particular, the results provide support for both Graded Salience (Giora, 1997) and Relevance Theory (Sperber & Wilson, 1995) frameworks for the effect of sarcasm on discourse processing by noting that comprehenders, given different discourse contexts, can selectively utilize Sarcasm in a context-dependent manner.

REFERENCES

Allopenna, P., Magnuson, J. & Tanenhaus, M. (1998). Tracking the time course of spoken word recognition using eye-movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.

Altmann, G. (2004). Language-mediated eye-movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93,* B79-B87.

Baayen, R. H. (2008). Analyzing Linguistic Data. A Practical Introduction to Statistics Using R. Cambridge University Press.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Balota, D. & Lorch, R. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12(3)*, 336-345.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B. & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445-459.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R. & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger Syndrome or hjigh-functioning Autism. *Journal of Autism and Developmental Disorders*, *29(5)*, 1999.

Bates, D., Maechler, M. & Bolker, B. (2011). lme4: linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4 (last visited 8/23/2012).

Blasko, G. D. & Connine, C. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning Memory, & Cognition*, *19(2)*, 295-308.

Blutner, R. & Sommer, R. (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language*, *27*, 359-367.

Braun, B. & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, *25*, 1024-1043.

Britt, A. & Mirman, D. (June 2012). *Robustness to individual differences: A comparison of bin-by-bin t-tests and growth curve analyses* (Language & Cognitive Dynamics Laboratory Technical Report 2012.02).

Bryant, G. A. & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and Symbol*, *17(2)*, 99-117.

Capelli, C., Nakagawa, N. & Madden, C. (1990). How children understand sarcasm: The role of context and intonation. *Child Development, 61(6)*, 1824-1841.

Cheang, H. & Pell, M. (2008). The sound of sarcasm. *Speech Communication, 50*, 366-381.

Chen, L. & Boland, J. E. (2008). Dominance and context effects on activation of alternative homophone meanings. *Memory & Cognition*, *36(7)*, 1306-1323.

Chevallier, C., Noveck, I., Happé, F. & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, *49*, 507-517.

Colston, H. L. & Lee, S. Y. (2004). Gender differences in verbal irony use. *Metaphor & Symbol*, *19(4)*, 289-306.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84-107.Cutler, A., Dahan, D. & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40(2)*, 141-201.

Dahan, D., Tanenhaus, M. K. & Chambers, C. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*, 292-314.

de Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *Journal of Verbal Learning & Verbal Behavior, 22,* 417-436.

de Groot, A. M. B. (1984). Primed lexical decision: Combined effects of the proportion of related prime-target pairs and the stimulus-onset asynchrony of prime and target. *Quarterly Journal of Experimental Psychology, 36A*, 253-280.

Dews, S. & Winner, E. (1995). Muting the meaning: A social function of irony. *Metaphor and Symbolic Activity*, *10(1)*, 3-19.

Duffy, S. A., Morris, R. K. & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory & Language*, *27*, 429-466.

Gibbs, R. W. (1980). Spilling the bean on understanding and memory for idioms in conversation. *Memory & Cognition*, *8*, 449-456.

Gibbs, R. (1984). Literal meaning and psychological theory. *Cognitive Science*, *8*, 275-304.

Gibbs, R. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General, 115(1)*, 3-15.

Gibbs, R. (1989). Understanding and literal meaning. *Cognitive Science*, *13*, 243-251.

Gibbs, R. & O'Brien, J. (1991). Psychological aspects of irony understanding. *Journal of Pragmatics*, *16*, 523-550.

Gibbs, R. (1994). The poetics of mind: Figurative thought, language, and understanding. Cambridge, MA: Cambridge University Press.

Gibbs, R. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, *34*, 457-586.

Giroa, R. (1995). On irony and negation. *Discourse Processes*, *19*, 239-264,

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *7*, 183-206.

Grainger, J., Van Kang, M. N. & Segui, J. (2001). Cross-modal repetition priming of heterographic homophones. *Memory & Cognition*, *29(1)*, 53-61.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), Syntax and semantics III: Speech acts (pp. 41-58). New York: Academic Press.

Griffin, Z. M. (1999). Frequency of meaning use for ambiguous and unambiguous words. *Behavior Research Methods, Instruments, & Computers*, *31(3)*, 520-530.

Happé, F. (1995). Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbolic Activity*, *10*, 275-295.

Hogaboam, T. W. & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, *14*, 265-274.

Huettig, F. & Altmann, G. (2005). Word meaning and the control of eye fixation: semantic competitor effects in the visual world paradigm. *Cognition, 96*, B23-B32.

Huettig, F. & Altmann, G. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15(8),* 985-1018.

Huettig, F., Quinlan, P., McDonald, S. & Altmann, G. (2006). Models of high-dimensional semantic space predict language-mediated eye-movements in the visual world. *Acta Psychologica, 121*, 65-80.

Hutchison, K. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10(4)*, 785-813.

Ivanko, S. L., Pexman, P. M. & Olineck, K. M. (2004). How sarcastic are you? Individual differences and verbal irony. *Journal of Language and Social Psychology*, *23(3)*, 244-271.

Joordens, S. & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(5)*, 1083-1105.

Jorgensen, J., Miller, G. & Sperber, D. (1984). Test of the mention theory of irony. *Journal of Experimental Psychology: General*, *113(1)*, 112-120.

Kerswell, L., Siakaluk, P. D., Pexman, P. M., Sears, C. R. & Owen, W. J. (2007). Homophone effects in visual word recognition depend on homophone type and task demands. *Canadian Journal of Experimental Psychology*, *61(4)*, 322-327.

Klin, C. & Drumm, A. (2010). Seeing what they read and hearing what they say: Readers' representation of the story characters' world. *Psychonomic Bulletin & Review*, *17(2)*, 231-236.

Kreuz, R. (1987). The subjective familiarity of English homophones. *Memory & Cognition*, *15(2)*, 154-168.

Kreuz, R. & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General, 118(4)*, 374-386.

Kucera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Lucas, M. (1987). Frequency effects on the processing of ambiguous words in sentence contexts. *Language and Speech*, *30(1)*, 25-46,

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 6,* 618-630.

MacDonald, M. C., Pearlmutter, N. J. & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.

Mason, R. & Just, M. (2007). Lexical ambiguity in sentence comprehension. *Brain Research*, *1146*, 115-127.

McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99(3)*, 440-466.

McNamara, T. P. & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory & Language, 27*, 545-559.

McQueen, J. & Viebahn, M. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology, 60(5)*, 661-671.

McRae, K. & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24,* 558-572.

Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Meyer, D. E. & Schvaneveldt, R. W. & Ruddy, M. G. (1972, November). *Activation of lexical memory*. Paper presented at the 13th Annual Meeting of the Psychonomic Society, St. Louis, MO.

Mirman, D., Dixon, J. & Magnuson, J. (2008a). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*, 475-494.

Mirman, D., Magnuson, J. S., Strauss, T. J. & Dixon, J. A. (2008b). Effects of global context on homophone ambiguity resolution. In *Proceedings of the 30th annual meeting of the cognitive science society*, July 2008, Washington, D. C.

Miyake, A., Just, M. A. & Carpenter, P. A. (1994) Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neural contexts. *Journal of Memory & Language*, *33*, 175-202.

Moss, H. E., Hare, M. L., Day, P. & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science, 6,* 413-427.

Moss, H. E., McCormick, S. F. & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes*, *12(5/6)*, 695-731.

Nakassis, C. & Snedeker, J., 2002. Beyond sarcasm: Intonation and context as relational cues in children's recognition of irony. In: Greenhill, A., Hughs, M., Littlefield, H., Walsh, H. (Eds.), Proceedings of the Twenty-Sixth Boston University Conference on Language Development. Cascadilla Press, Somerville, MA. pp. 429-440.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Neill, W. T., Hilliard, D. V. & Cooper, E. (1988). The detection of lexical ambiguity: Evidence for context-sensitive parallel access. *Journal of Memory & Language*, *27*, 279-287.

Nelson, D. L., McEvoy, C. L. & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://usf.edu/FreeAssociation/.

Perea, M. & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition*, *62*, 223-240.

Perea, M. & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological research, 66,* 180-194.

Sarcasm: Do you hear it now?. Peters, S., Wilson, K., & Almor, A. (2013). The 26th Annual CUNY conference on sentence processing. Poster. Columbia, SC.

Pexman, P. M. & Olineck, K. M. (2002). Does sarcasm always sting? Investigating the impact of ironic insults and ironic compliments. *Discourse Processes*, *33(3)*, 199-217.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–226.

Posner, M. & Snyder, C. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191-201.

Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. *Journal of Psycholinguistic Research, 29(5)*, 483-495.

Salverda, A. P., Brown, M. & Tanenhaus, M. K. (2011). A goal-based perspective on eye-movements in visual-world studies. *Acta Psychologica*, *137(2)*, 172-180.

Seidenberg, M., Tanenhaus, M., Lieman, J. & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14(4)*, 489-537.

Shelton, J. R. & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1191-2010.

Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition.* Cambridge, MA: Harvard University Press.

Sperber, D. & Wilson, D. (1995). *Relevance: Communication and Cognition: 2nd Edition.* Cambridge, MA: Blackwell Publishers Inc.

SR Research, Ltd. (2013). Experiment Builder & Data Viewer Software (Versions 1.10.165 & 1.11.1). Available from http://www.sr-research.com/.

Starkweather, J. (2010). Linear Mixed Effects Modeling using R. Unpublished Manuscript.

Sullivan, K., Winner, E. & Hopfield, N. (1995). How children tell a lie from a joke: The role of second-order mental state attributions. *British Journal of Developmental Psychology*, *13*, 191-204.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645-659.

Tanenhaus, M., Magnuson, J., Dahan, D. & Chambers, C. (2000). Eye-movements and lexical access in spoken-language comprehension: evaluating a linking hypothesis

between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29(6)*, 557-580.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science, 268*, 1632-1634.

Weingartner, K. & Klin, C. (2005). Perspective taking during reading: An on-line investigation of the illusory transparency of intention. *Memory & Cognition, 33(1)*, 48-58.

Weingartner, K. & Klin, C. (2009). Who knows what? Maintaining multiple perspectives during reading. *Scientific Studies of Reading, 13(4)*, 275-294.

Wilson, D. & Wharton, T. (2006). Relevance & Prosody. *Journal of Pragmatics*, *38*, 1559-1579.

Yee, E. & Sedivy, J. (2006). Eye-movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(1)*, 1-14.

Yee, E., Overton, E. & Thompson-Schill, S. L. (2009). Looking for meaning: Eye-movements are sensitive to overlapping semantic features, not association. *Psychonomic Bulletin & Review, 16(5)*, 869-874.

APPENDIX A: HOMOPHONE INFORMATION

Table A.1. Homophone information.

#	Ambiguous Ortho-graphy in Experiment	Item	Associate	Neutral	Item (Prime) Freq.	Associate (Target) Freq.	Neutral Baseline Freq.	Social Contextua l Ratings (1-5) M (*sd*)
1	N	Beach	Sand	Jump	61	28	24	3.10 (0.27)
1	N	Beech	Bark	Depot	6	14	13	1.83 (0.21)
2	N	Base	Home	Found	91	547	536	2.33 (0.22)
2	N	Bass	Guitar	Label	16	19	19	2.63 (0.26)
3	N	Dock	Pier	Axle	8	3	5	2.73 (0.25)
3	N	Doctor	Nurse	Essay	100	17	19	3.60 (0.24)
4	N	Fir	Tree	Roof	2	59	54	1.97 (0.23)
4	N	Fur	Coat	Pilot	13	43	44	3.20 (0.24)
5	N	Flour	Cake	Dentist	8	13	12	3.00 (0.29)
5	N	Flower	Rose	King	23	86	88	3.03 (0.26)
6	N	Fairy	Wand	Cube	6	1	1	2.30 (0.25)
6	N	Ferry	Boat	Phase	11	72	72	2.43 (0.28)
7	N	Hair	Brush	Journal	148	44	44	3.40 (0.26)

7	N	Hare	Rabbit	Atlas	1	11	12	2.43 (0.27)
8	N	Knight	Armor	Wharf	18	4	4	2.53 (0.26)
8	N	Night	Day	Came	411	686	622	3.57 (0.24)
9	N	Oar	Paddle	Lacey	0	1	2	2.36 (0.26)
9	N	Ore	Iron	Mold	3	43	45	2.20 (0.36)
10	N	Genes	Body	Door	9	276	312	3.67 (0.19)
10	N	Jeans	Pants	Deed	1	9	8	3.56 (0.27)
11	N	Sail	Ocean	Atom	12	34	37	2.77 (0.29)
11	N	Sale	Clothes	Cover	44	89	88	3.63 (0.23)
12	N	Son	Daughter	Newspaper	166	72	65	3.43 (0.26)
12	N	Sun	Moon	Risk	112	60	59	3.53 (0.25)
13	N	Suite	Hotel	Below	27	126	145	2.90 (0.26)
13	N	Sweet	Candy	Kick	70	16	16	3.70 (0.26)
14	N	Tea	Coffee	Judge	28	78	77	3.36 (0.29)
14	N	Tee	Club	Wall	5	145	160	2.20 (0.25)
15	N	Waist	Belt	Jet	11	29	29	3.86 (0.17)
15	N	Waste	Trash	Dame	35	7	7	3.40 (0.21)
16	Y	Trunk	Car	Kind	8	274	313	2.90 (0.23)
16	Y	Trunk	Elephant	Paramount	8	7	9	2.07 (0.21)
17	Y	Nut	Bolt	Tunnel	15	10	10	1.97 (0.21)

17	Y	Nut	Squirrel	Umpire	15	1	1	2.47 (0.24)
18	Y	Bow	Arrow	Thunder	15	14	14	2.53 (0.24)
18	Y	Bow	Ribbon	Mineral	15	12	12	2.53 (0.24)
19	Y	Bulb	Lamp	Mist	7	18	14	2.97 (0.27)
19	Y	Bulb	Tulip	Isle	7	4	5	1.87 (0.19)
20	Y	Pipe	Smoke	Piano	20	41	38	2.80 (0.23)
20	Y	Pipe	Water	Left	20	442	480	2.40 (0.25)
21	Y	Diamond	Field	Period	8	274	265	1.83 (0.19)
21	Y	Diamond	Ring	Missile	8	47	48	3.73 (0.25)
22	Y	Bank	Money	Cannot	83	265	258	3.57 (0.23)
22	Y	Bank	Rock	Pain	83	75	74	1.97 (0.22)
23	Y	Ball	Dance	Shape	110	90	85	2.27 (0.23)
23	Y	Ball	Game	Stop	110	123	120	3.30 (0.23)
24	Y	Plant	Factory	Jacket	125	32	33	2.23 (0.17)
24	Y	Plant	Seed	Core	125	41	37	2.90 (0.24)
25	Y	Port	Ship	Edge	21	83	78	2.17 (0.24)
25	Y	Port	Wine	Rock	21	72	75	2.03 (0.22)
26	Y	Boxer	Gloves	Kitty	1	7	7	2.27 (0.22)
26	Y	Boxer	Shorts	Pages	1	29	31	3.13 (0.23)
27	Y	Batter	Baseball	Senate	2	57	62	2.43 (0.25)

27	Y	Batter	Dough	Leap	2	13	14	2.90 (0.26)
28	Y	Bat	Ball	Test	18	110	119	2.83 (0.28)
28	Y	Bat	Cave	Boom	18	9	8	2.23 (0.20)
29	Y	Crane	Bird	Keys	5	31	34	1.93 (0.23)
29	Y	Crane	Construct -ion	Significant	5	95	85	2.53 (0.29)
30	Y	Pit	Cherry	Kidney	14	6	6	1.90 (0.17)
30	Y	Pit	Hole	Pick	14	58	55	2.47 (0.22)

Neutral baseline was matched based on letter number and frequency in Kucera & Francis (1967) to the associates that could appear as a target to the homophone primes. Frequency measures are taken from Kucera & Francis (1967).

APPENDIX B: EXPERIMENT DISCOURSES

Table B.1. Experimental Discourses.

Experiment	Item	Speaker A	Speaker B
2	1	My favorite part is that beach over there.	This time of year, it's especially pretty.
2	2	We can't start until someone gets the bases out of storage.	Don't worry, it will be so much fun.
2	3	It sounds like it's the perfect time for a trip to the dock this afternoon.	Yeah, I haven't been there in so long.
2	4	I'm not sure you can overlook the fur when you make a decision.	How do I decide, they all look so great.
2	5	Just make sure you remember to pick up the flower for me.	I suppose you need it in order for the plan to be carried out.
2	6	We can't change the size of the ferry in this case.	Well I think it is definitely big enough.
2	7	It's a pretty big competition, I hope my hare is okay.	I think it will be fine, and over quickly.
2	8	If you are afraid of the night it won't work.	I think it's absolutely perfect for the plan.
2	9	I would feel better if we had found the oar we were looking for.	Well, we could just prepare better for next time.
2	10	She has the best jeans in the group.	I can't imagine why you would think that.
2	11	It's a very unique sale for sure.	I think that it will be very effective.
2	12	Having such a bright sun makes me happy.	You really lucked out on that one.
2	13	I always prefer to get a suite if I can.	It's always nice to spoil yourself.
2	14	I can't find the tea anywhere.	We need it, so hopefully you can find it.
2	15	I need to reduce my waste significantly this year.	When you decide how to do it, let me know how it works out.
2	16	It has a pretty huge trunk, most definitely.	I'm not sure it's the biggest one I've ever seen.
2	17	Hand me the bag of nuts please.	I thought you had everything you needed.
2	18	Someone needs to tighten the bow before we start.	It does look a little loose.
2	19	I feel like I have to buy bulbs every year.	Maybe they're more delicate than you realized.

2	20	Just because it’s so old doesn’t mean the pipe won't work.	You'll have to let me know if lasts through the holidays.
2	21	I wouldn't hang on to an old diamond just to keep it in the family.	If I sell it, I'm sure to get a good return.
2	22	I’d never miss a trip down to the bank with you.	Whenever I go there, I always have some excitement.
2	23	I heard everyone talking about the ball again.	Yes, it is the most exciting thing this year.
2	24	I can’t believe there is so much fuss over a new plant already.	The prospect of it is just thrilling.
2	25	You seem to like the port here.	This is the best one yet.
2	26	You've never had this much trouble picking boxers before today.	I've narrowed it down, but the choice is overwhelming.
2	27	I think this is the worst batter ever.	I think there has probably been worse.
2	28	Be careful with that bat please.	Yeah that's really dangerous.
2	29	I can’t wait to look outside and see a crane across the street.	I think the project will add a lot to the neighborhood.
2	30	I’d probably enjoy it more if not for the pit to deal with.	You seem to be enjoying it anyway.
3	1	It’s very pretty, especially this time of year.	Well, my favorite part is the beach over there.
3	2	This is going to be so much fun.	As long as someone remembers to get the bases out of storage.
3	3	I haven’t been there in so long.	It sounds like the perfect time for a trip to the dock this afternoon.
3	4	It’s so hard to make a decision.	I’m not sure you can overlook the fur when you make a decision.
3	5	We’re almost done with everything for the party.	Just make sure you don’t forget to pick up the flower for me.
3	6	The sign needs to be big enough for everyone to see.	We can't change the size of the ferry in this case.
3	7	Thank you for coming with me.	Well, it's a pretty big competition for you and your hair today.
3	8	I think the setting is absolutely perfect for the plan.	You would need to be a fan of the night for it to work.
3	9	You’re a good friend for talking me into this.	I wish we had found the oar we were looking for.
3	10	She has really good luck.	You can’t blame her for getting the best jeans in the group.
3	11	That is going to get a lot of attention.	It's a very unique sale for sure.
3	12	You really lucked out today.	Yeah, having such a bright sun makes me happy.
3	13	It's always nice to spoil yourself once in a while.	Yes, I always prefer to get a suite once in a while.
3	14	Hopefully I'll be able to find it.	Please tell me you didn’t leave the tea back there.
3	15	I’m trying to cut down for the new year, that's all.	If you succeed in reducing your waste, let me know.
3	16	I'm not sure if it's the biggest one	It's a pretty massive trunk in my opinion.

		I've ever seen.	
3	17	I think you have everything you need.	Something tells me we are missing a bag of nuts to add in.
3	18	I'm not sure if that's right.	Someone needs to tighten the bow before we're finished.
3	19	I feel like I have to buy these every year.	Maybe you just don't know how to take care of a bulb.
3	20	I haven't seen something like that in ages.	Just because it's so old doesn't mean its not a functional pipe still.
3	21	If I sell it, I may get a good return.	I don't think you need to hang on to an old diamond just to have it.
3	22	I'm glad you are going down there with me tomorrow.	I'd never miss a trip to the bank with you.
3	23	I heard everyone talking about it.	Yes, I don't understand why everyone is talking about one ball when we have so many.
3	24	I've seen a picture, it's going to look great.	I can't believe the fuss over a new plant, especially here.
3	25	I like this one a lot.	This is the best port yet.
3	26	This is not how I want to spend my Saturday.	I've narrowed it down but the choice boxers is overwhelming.
3	27	This is an interesting turn of events.	Well I'm sure there's been a worse batter at some point.
3	28	That looks really dangerous.	Don't worry, I'll be careful with the bat here.
3	29	I can't wait to look outside and enjoy the view.	Yes, I can't wait to look outside and see a crane across the street.
3	30	You seem to be enjoying yourself.	I'd probably enjoy it more if not for the pit to watch out for.

APPENDIX C: SPECTROGRAMS

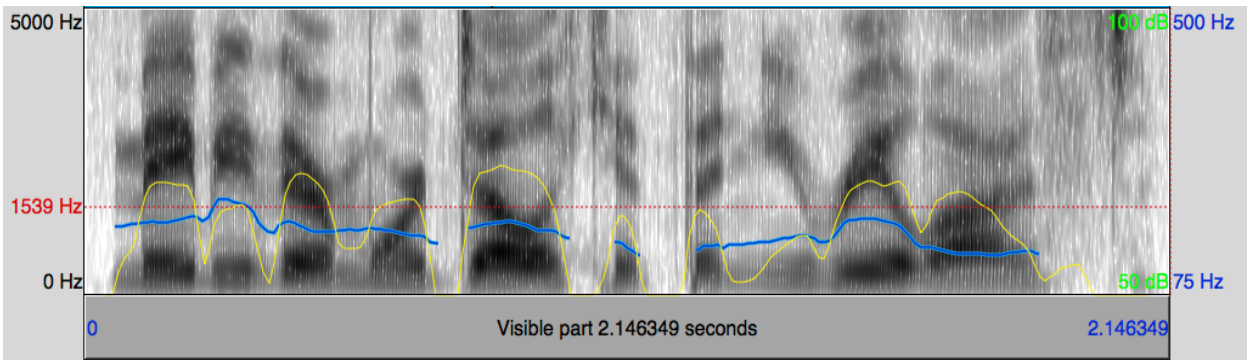

Figure C.1. Spectrogram of Experiment 2, Sentence 2, Sincere, phrase "Maybe they require more care than you realized."

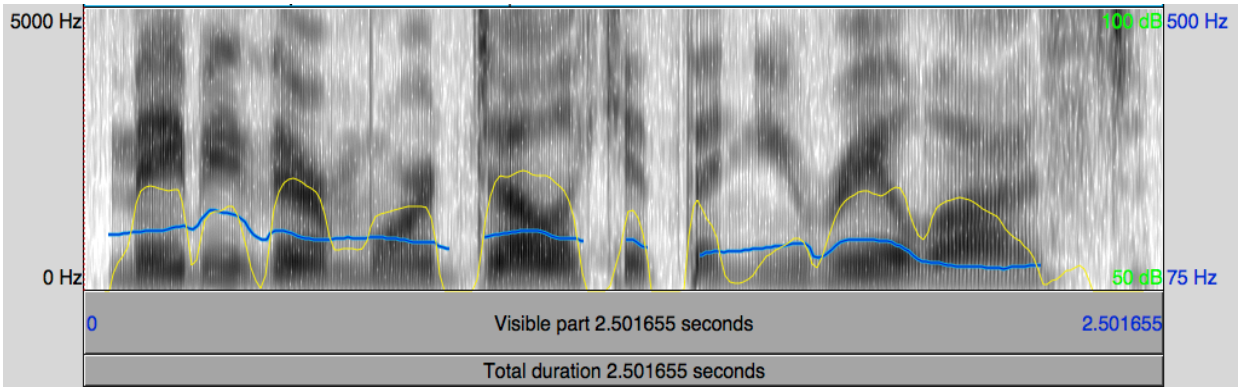

Figure C.2. Spectrogram of Experiment 2, Sentence 2, Sarcastic, phrase "Maybe they require more care than you realized."

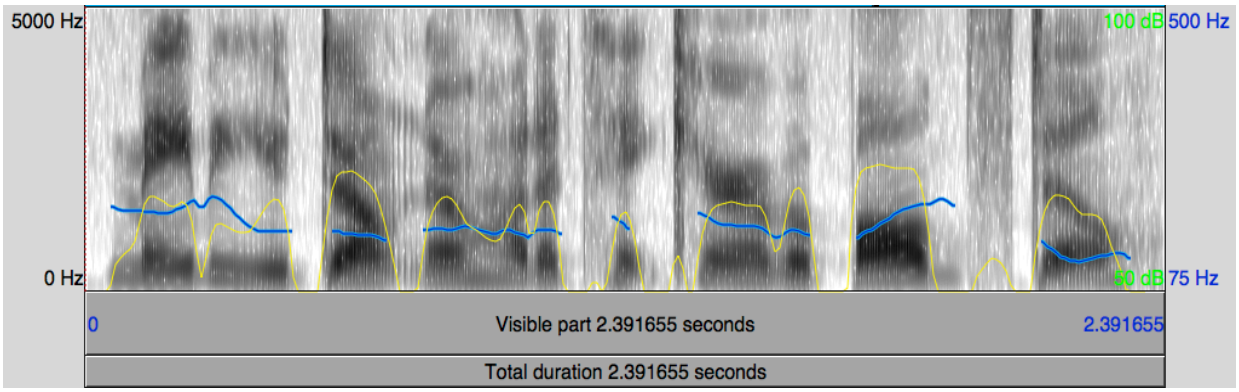

Figure C.3. Spectrogram of Experiment 3, Sentence 2, Sincere, phrase "Maybe you just don't know how to take care of a bulb properly."

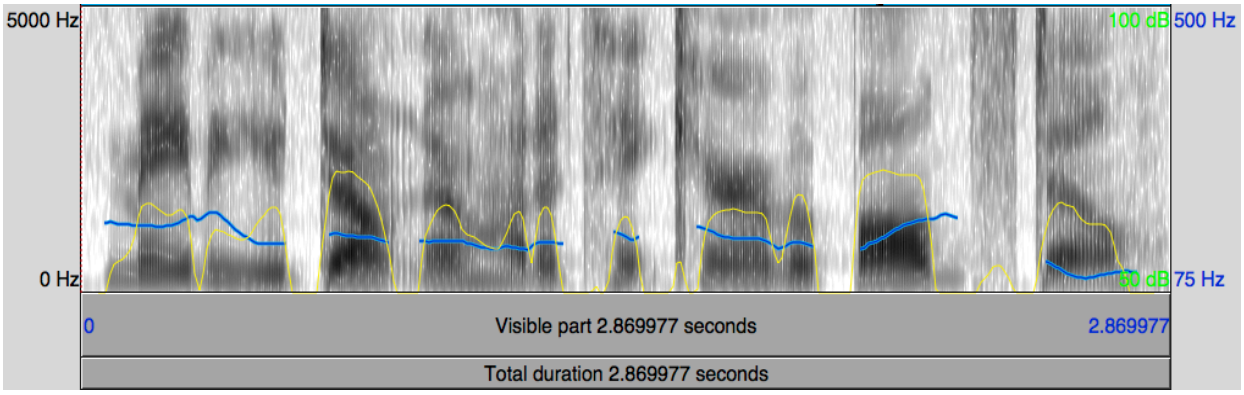

Figure C.4. Spectrogram of Experiment 3, Sentence 2, Sarcastic, phrase "Maybe you just don't know how to take care of a bulb properly."