How Second Language English Learners Interpret Sarcasm in English: A Survey

Jack Dolan

Irony is prevalent in most daily situations. Native English speakers tend to have trouble understanding whether a certain phrase or utterance was meant to be sarcastic. This article examines how second language (L2) English learners respond to certain questions meant to convey either sarcasm or sincerity. A survey is used and data analysis created to understand whether or not L2 English learners are on the same level as native English speakers in interpreting prosodic cues for sarcasm.

Introduction

While having dinner at a local restaurant with a colleague, Robert inadvertently spills mustard on his shirt. He is unable to remove the stain, and he returns home afterward to his family. As Robert walks in the door, his wife notices the stain and says, "You must've loved that dinner since you decided to bring it home with you," finishing the quip with a smrik on her face. Most native English speakers would chuckle at this, understanding that the utterance made by Robert's wife was ironic. However, it's quite possible that non-native English speakers would have trouble interpreting the irony of a situation.

When defining irony in language, Sperber et al. state that "an ironical utterance is traditionally analyzed as literally saying one thing and figuratively meaning another" (Sperber et al., 295). Jorgensen et al. agree that "an ironist uses a figurative meaning opposite to the literal meaning of the utterance" (Jorgensen, et al., 112). What might affect a listener's understanding of the irony used in certain situations? Clark and Gerrig discuss the interaction between the illocutionary and perlocutionary acts and assert that "a listener's understanding of an ironic utterance depends crucially on the common ground he or she believes is shared by the ironist and the audience—their mutual beliefs, mutual knowledge, and mutual suppositions" (Clark, et al., 124). Rockwell (2001) finds that intonations and speech patterns, or prosody, are also crucial for the understanding of ironic utterances (2001).

Cheang et al. state that "speech prosody (i.e., intonation and stress patterns) conveys many types of information to listeners" (Cheang, et al., 1394). Wilson and Wharton stated that prosodic features "typically create impressions, convey information about emotions or attitudes, or alter the salience of linguistically-possible interpretations rather than conveying distinct propositions or concepts in their own right" (Wilson, et al., 1559). Kreuz and Roberts and Cheang and Pell find that one subtype of irony in which specific interest prosody seems to play a very significant role is that of sarcasm.

Kreuz and Glucksberg separate sarcastic irony (sarcasm) from verbal irony by stating that "sarcastic irony involves the use of counterfactual statements to express disapproval, usually with intent to hurt or wound someone or some group of people," where-

as "in verbal irony a speaker expresses an attitude toward some object, event, or person by saying something that is not literally true" (Kreuz, et al., 374). Cheang and Pell note "the unique feature of sarcasm as a form of verbal irony is that it is chiefly used to express negative critical attitudes" (Cheang, et al., 1394). Certain prosodic indicators allow these negative expressions to be perceived as sarcasm. Indicators of sarcasm in language have been a topic of study largely within the last twenty years. Studies have shown that some facial cues are indicative of sarcasm (Rockwell, 2001). Rockwell found that when using sarcasm, speakers tend to use a lower pitch, slower speed, and higher volume than in regular speech (Rockwell, 2000). Cheang and Pell also found that in addition to a slower rate of speech, the fundamental frequency of a speaker's voice is usually lower when indicating sarcasm (Cheang, et al, 2008).

Much work has also been done on the acoustical features of sarcasm cross-linguistically. For example, Cheang and Pell did a study on the acoustic markers of Cantonese and English (2008). Upon analysis, Cheang and Pell found that, although etymologically and culturally different, English and Cantonese employed a common prosodic cue when expressing sarcasm: slower speech rate. However, Cantonese has a higher fundamental frequency when expressing sarcasm, as opposed to English where there is a lower fundamental frequency.

In another study, Capelli found that children under the age of twelve didn't understand sarcastic utterances without the specific use of sarcastic intonation, even when provided with contextual clues (Capelli, et al, 1990). Gibbs has also found that context affects whether or not the listener understands the ironic utterances (1994). Bryant and Fox Tree (2002) found that prosodical cues produced in spontaneous language (i.e., language not produced by professional voice actors) help listeners to discern between sarcastic and sincere speech (2002).

As can be seen in these references, many studies have been conducted on the acoustical features of sarcasm cross-linguistically and on the role that context can play in the interpretation of certain sarcastic and ironic utterances. However, very little research has been dedicated to the study of L2 English learners' understanding of sarcastic and sincere utterances based on con-

text and prosodical cues. This study examines L2 English learners' ability to pick up on verbal versus written sarcasm, as well as their ability to distinguish between sincere and sarcastic utterances, using written context clues and spoken prosodical cues.

The two hypotheses in this study are: 1) it will be more difficult for L2 English learners to pick up on orthographic sarcasm vs. verbal sarcasm, compared to native English speakers; and 2) native speakers will be more accurate in rating sarcastic and sincere utterances.

The implications of this study are numerous. The Bureau of Labor Statistics (2012) estimates that in the US, the number of English language teachers will increase by 12.2% by the year 2022, brining in a total of 23,600 new jobs. However, the biggest difficulty of teaching English is the ability to learn all of the acoustic cues needed in conversation, such as raising intonations for questions and knowing where to put the stress to convey meaning. Understanding this aspect of language will help language teachers know how to better assist L2 English learners with comprehension of sarcasm in a variety of situations including the classroom, workplace, and other social situations.

Methodology

Participants

In this study, fifteen L2 learners of English and fifteen native English speakers participated. Means of obtaining these participants included email and social media posts. Participants were of various ethnicities, ranging from Koreans and Peruvians to Hondurans and Mauritians, as well as of various ages and sex. Below is a table providing a summary of the participants' information:

Nationality	Native Language	Average # Years of Formal English Study	Average Age	Male	Female
American	English	15.13	33.27	10	5
Guatemalan	Spanish	2	40	1	0
Honduran	Spanish	6.5	42	0	2
Peruvian	Spanish	2	22	1	0

Nationality	Native Language	Average # Years of Formal English Study	Average Age	Male	Female
Mexican	Spanish	1.5	1.5	1	1
Brazilian	Portuguese	4	22	0	1
Mauritians	Mauritian Creole/French	10	17	1	0
South Korean	Korean	8.4	22	0	5
Chinese	Mandrin	8	22	0	1
Chinese	Cantonese	21	22	1	0

Instrument

A survey consisting of ten questions—five written phrases and five audio files—was utilized for this study. Each written phrase had a preceding scenario to help the reader clearly see whether sincerity or sarcasm was being conveyed. One example is given below:

Figure 1: *Written survey sample*

Read the following statement:

(Jane is Dan's wife. Dan walks in the door completely drenched from the rain storm outside.)

Jane: Nice weather out today, huh, honey?

	Sincere	Somewhat Sincere	Somewhat Sarcastic	Sarcastic
Jane's statement was:	○	○	○	○

Each audio file contained an utterance in which the intonation conveys either sarcasm or sincerity. It should be noted that no context was given to help the listeners' interpretation so as to test the ability of the listeners to detect prosodical cues. An example is provided below:

Figure 2: *Audio survey sample*

Listen to the following sound file:

	Sincere	Somewhat Sincere	Somewhat Sarcastic	Sarcastic
The utterance sounded:	○	○	○	○

Responses were recorded using a Likert scale for each utterance. Choices ranged from 1 to 4, 1 being "Sincere," 2 being "Somewhat Sincere," 3 being "Somewhat Sarcastic," and 4 being "Sarcastic." Lack of a neutral number was designed to force the participants to make a decision based on what they heard. Each question was untimed in order to allow the listener to replay the sound files as much as needed and to feel comfortable with his or her decision.

Each survey was taken electronically via Qualtrics.com. Surveys were taken primarily on laptop and desktop computers. No compensation was provided to those who participated, and each participant decided to participate under their own volition.

Data Analysis

The Qualtrics database recorded each response and provided the ability to export the data to data analysis software. Microsoft Excel was utilized for the analysis of the relationships between the responses recorded and the demographics solicited in the survey.

The data was divided into two categories based on whether the speaker was from the United States and whether the speaker was a native English speaker. Means were computed for each group (native and non-native English speakers) and measured based on individual question and type, overall average for the sarcastic and sincere question sets, and written vs. spoken question type.

Various bar graphs were generated to show graphic representations of the data. *T*-tests of independent samples were used to show whether the differences among the various data sets were significant or not and were represented in the bar graphs.

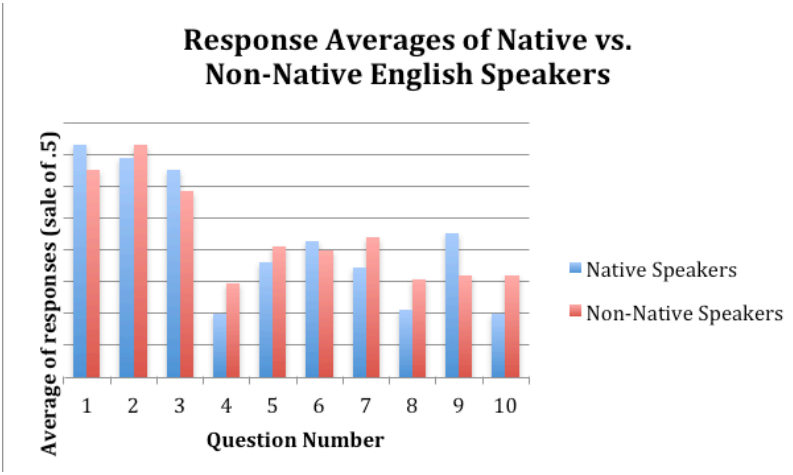
Scatter plots were created to show the correlation between the average of the sarcastic responses and the factors such as the amount of time spent formally studying English and the participants' age and gender. Correlations were computed in Excel and were interpreted using the following guidelines: 0.10 (Weak), 0.25 (Moderate), and 0.50 (Strong).

Results

The averages for all of the questions were calculated and are shown in Figure 3. The data for each item is divided into two

groups—native and non-native English speakers. The means for all items are calculated on a Likert scale of 1–4; 1 represents complete sincerity and 4 represents complete sarcasm.

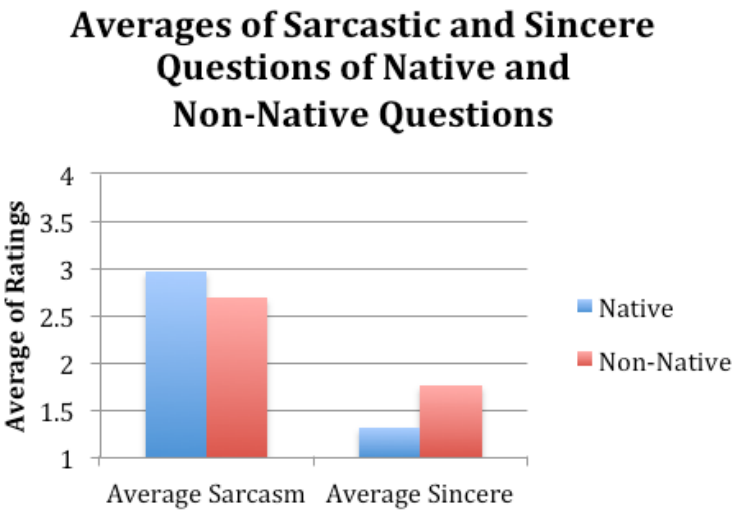
Figure 3: *Answer averages per native and non-native speakers*



Items 1, 2, 3, 6, and 9 were meant to convey sarcasm. Items 4, 5, 7, 8, and 10 were meant to convey sincerity. The trend in the data shows that the native speakers were closer to the original expectations for each question than the non-native speakers 90% of the time. Question 2, however, shows a different result. Question 2 was meant to convey sarcasm. The non-native speakers' average understanding of sarcasm was closer to 4 (M = 3.67, SD = 0.90) than the average of native speakers (M = 3.47, SD = 0.52).

Figure 4 represents the average of all sarcastic and sincere questions per group. These averages were calculated by combining each participant's answers for all of the sincere and sarcastic questions, then by averaging those means. The results show that, with sarcastic questions, native speakers were able to rate closer to the intended result (a score of 4) than non-native speakers. The same pattern was found with the sincere questions, showing that native speakers tended to rate closer to the intended result (a score of 1) than non-native speakers.

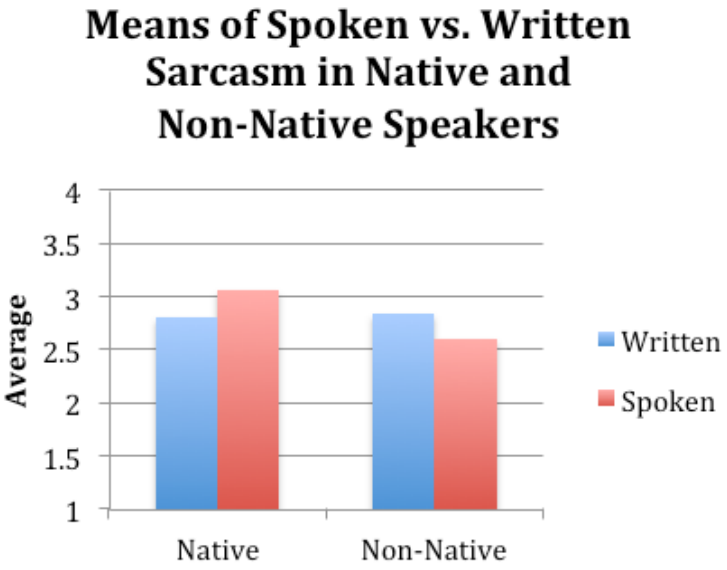
Figure 4: *Sincere vs. Sarcastic Question Averages*



Two *t*-tests were run to analyze whether the difference in averages between native and non-native speakers' answers were significant, with the alpha being $p < .05$. The first *t*-test measured the difference between native and non-native speaker averages for sincere questions. The results of the first *t*-test ($p = .011$) revealed a significant difference between the two groups. This test showed that non-native speakers tended to rate sincere questions as more sarcastic than native speakers. The second *t*-test measured the relationship between the averages of native and non-native speakers for sarcastic questions. The results of the *t*-test ($p = .066$) show that the difference between the groups' average responses to the sarcastic items was not significant.

To test for result accuracy between spoken versus written sarcastic responses, the averages for native and non-native speakers were calculated for each individual speaker. A mean was then calculated from those averages. The means were divided into two groups: native English speakers and non-native English speakers. These results were gathered to test the hypothesis that native speakers tend to rate written sarcasm lower than that of spoken sarcasm. The results are presented in Figure 5.

Figure 5: *Average Spoken vs. Written Sarcasm Responses*



As can be seen in the data, non-native speakers rated written sarcastic responses at a higher average (M = 2.83, SD = 0.59) than native speakers (M = 2.8, SD = 0.53). However, native speakers rated spoken sarcastic items as being much more sarcastic (M = 3.07, SD = 0.47) than did non-native speakers (M = 2.60, SD = 0.44).

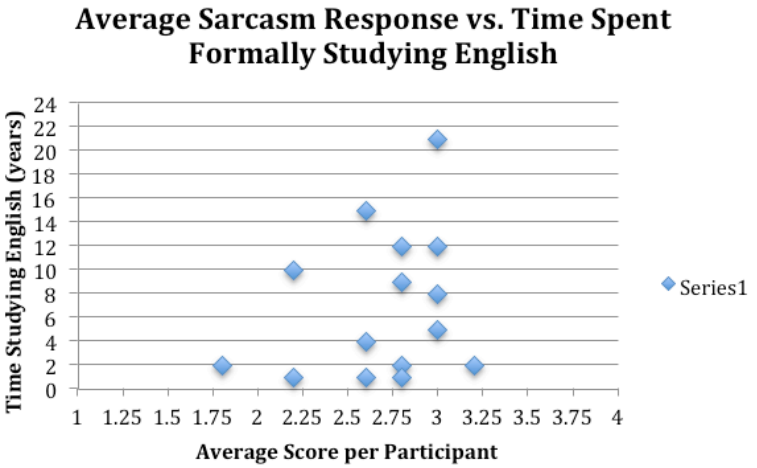
More *t*-tests were conducted to find whether the differences in these averages were significant. Again, the alpha was set at $p < .05$. The third *t*-test measured the differences between the natives' responses to spoken and written sarcasm. These responses had a result of p = 0.157, showing that the difference between the responses for written versus spoken sarcasm wasn't significant. A fourth *t*-test was conducted that measured the differences between the non-natives' responses to spoken and written sarcasm. These had a result of p = 0.229, also showing that the difference between non-native responses for written sarcasm and spoken sarcasm was also not significant.

A fifth *t*-test measured the differences between the responses of natives and non-natives to spoken sarcasm. This had a result

of p = 0.001. This difference suggests that the native speakers (as seen in Figure 5) were more capable of identifying the prosodical cues of sarcasm in the spoken utterances than the non-native speakers were. A sixth *t*-test measured the differences between the responses by natives and non-natives to written sarcasm. This had a result of p = 0.871. This shows that both native and non-native speakers tended to rate the written responses as sarcastic, with the difference between the two groups being insignificant.

Correlation tests were run between the averages for each non-native's response to sarcastic items and these factors: amount of formal English study, age, and gender. Figure 6 represents these correlations.

Figure 6: *Average score per participant for sarcastic utterances vs. time spent formally studying English*



Formal study was defined as time spent studying English in a classroom. Time spent in formal study was solicited from each participant and was rounded to the nearest year. The result was a positive correlation of 0.267, showing a weak relationship between the two variables.

Figure 7: *Age of participant vs. average score per participant for sarcastic utterances*

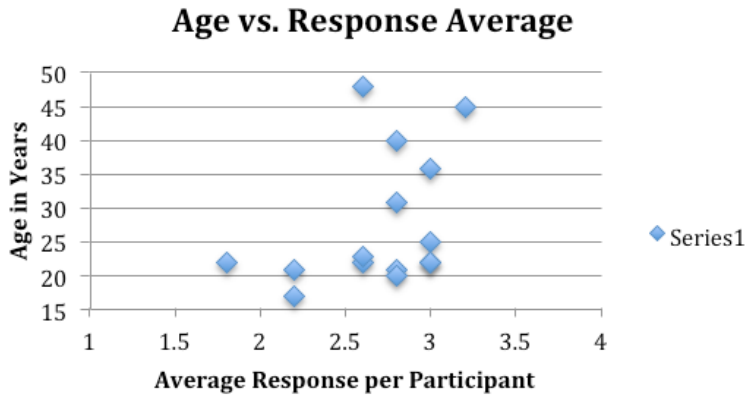


Figure 7 (above) represents the average score per participant for each sarcastic utterance and the age of each participant. A correlation test was run, resulting in a positive correlation of 0.381. This suggests a weak to moderate relationship between age and average response.

Conclusions

The purpose of this study was to measure 1) whether L2 English learners interpreted written sarcastic utterances as more sincere than spoken utterances and 2) whether L2 English learners or native English speakers were better able to interpret the intended conveyance of sincerity or sarcasm.

The first hypothesis proposed in the introduction was that the L2 participants would have a harder time interpreting written sarcasm vs. spoken sarcasm. That hypothesis proved to be false, and the reverse was true. Figure 5 shows that the average for written sarcasm was 2.83, and the average for spoken sarcasm was much lower at 2.60. This suggests that L2 learners tend to understand sarcasm better when given contextual information than when they are given only the prosody used in spoken utterances.

Figure 5 shows that the differences between the natives' and the non-natives' answers weren't significant, suggesting that their understanding of sarcasm is not very different. However, the

difference between a native and a non-native's understanding of spoken sarcasm in particular was very significant, suggesting that non-native speakers interpret sarcasm differently than do native speakers. This should be investigated further, focusing on factors that might affect understanding such as intonation patterns and societal factors.

Also in Figure 5, the differences between the natives' and the non-natives' responses to items conveying written sarcasm wasn't significant. This suggests that, when given context for clarification, both groups were able to interpret the sarcasm conveyed with almost the same accuracy.

From the data presented in Figure 4, it can be concluded that native speakers of English were able to interpret sarcasm and sincerity in English better than L2 speakers, which supports one of the hypotheses proposed at the beginning of this study. Other studies can be performed to analyze this reason more in-depth. For example, what factors affect this? Possible factors could be culture and country of origin. Several factors were investigated already in this study, specifically gender, amount of formal English study, and age. Results showed that there was a moderate positive relationship between all three factors with respect to average responses.

Figure 6 suggests that time spent formally studying English may slightly affect an L2 speaker's understanding of sarcastic and sincere prosodical cues. Those participants who tended to rate items conveying sarcasm as more sarcastic tended to have more years of formal English study. However, that was not always the case. More in-depth study should be conducted in this area to understand more accurately how much formal classes in English affect sarcasm interpretation by L2 English learners.

Figure 7 suggests that the age of the L2 English learner may also slightly affect their understanding of the previously mentioned prosodical cue. The older the participant, the more likely they were to correctly understand sarcasm. Future study should be dedicated to this, controlling such factors as time spent living in the U.S. This would more accurately show whether those speakers who were older and had spent more time living in the U.S. truly have an advantage over the younger speakers.

Figure 3 provides a basis for further studies into responses of L2 English learners to sarcasm. This study was limited by having

very few participants that were mostly related in some way to the researcher, as well as by having only one speaker in the sound samples. Figure 3 indicates that native speakers were usually more accurate in rating each question according to the original premise of the researcher. However, further study with more participants and more analysis could provide deeper analysis and application for this data.

Several other limitations should be considered for future study of this topic or replication of the experiment. First is the medium. Many of the participants took the quiz on their own personal computers and devices. Several complained about how the audio didn't work on their mobile devices, providing the need to have a controlled location where the survey would be taken. One suggestion would be a specific computer lab where there would be fewer distractions. Another limitation includes a lack of diversity in the participants. Most of the native participants had had previous interaction with the researcher. Since the researcher provided the sound samples, this could have provided an inaccurate basis on which to compare the results from the non-native speakers. Another important factor to use in future studies would be to test whether context truly played a significant role in the understanding of the native and non-native speakers. In this study, only the written questions provided context. Future studies should measure the extent to which context plays a significant role in enabling the participants to better detect sarcasm.

Overall, the study was successful in exploring the understanding of prosodical cues by non-native speakers of English. Sarcasm is a prevalent factor in the daily life of English speakers, ranging from social interaction with family members to stand-up comedy, and from TV sitcoms to board meetings.

The understanding of sarcasm and what makes something sarcastic has important implications for second language teaching. As the immigration of non-native English speakers to English-speaking countries continues, understanding and teaching the implications of sarcasm will help L2 speakers gain a more in-depth understanding of the sometimes unspoken rules of prosody in English.

It can be concluded from this study that 1) native English speakers tend to understand sarcasm and sincerity more than non-native speakers, and 2) several factors such as age and time

spent formally studying English may affect the understanding of an L2 English learner with regards to sarcasm and sincerity. The data indicates that there is a difference in the ability of native and non-native speakers' to understand sincerity and sarcasm. As such, an effort should be made to better understand why there is a gap and what can be done differently to bridge that gap.

Works Cited

Anolli, Luigi, Rita Ciceri, and Maria Giaele Infantino. "From 'Blame by Praise' to 'Praise by Blame': Analysis of Vocal Patterns in Ironic Communication." *International Journal of Psychology* 37, no. 5 (2002): 266-276.

Bureau of Labor Statistic. "Employment Projections" (2012). Retrieved from http://data.bls.gov/projections/occupationProj.

Bryant, Gregory A., and Jean E. Fox Tree. "Is There an Ironic Tone of Voice?" *Language and Speech* 48 (2005): 257–277.

Capelli, Carol A., Noreen Nakagawa, and Cary M. Madden. "How Children Understand Sarcasm: The Role of Context and Intonation." *Child Development* 61, no. 6 (1990): 1824–1841.

Cheang, Henry S., and Marc D. Pell. "Acoustic Markers of Sarcasm in Cantonese and English." *The Journal of the Acoustical Society of America* 126, no. 3 (2009): 1394–1405.

Cheang, Henry S., and Marc D. Pell. "The Sound of Sarcasm." *Speech Communication* 50, no. 5 (2008): 366–81.

Clark, Herbert H., and Richard J. Gerrig. "On the Pretense Theory of Irony." *Journal of Experimental Psychology: General* 113, no. 1 (1984): 121–26.

Gibbs, Raymond W. *The Poetics of Mind*. Cambridge: Cambridge University Press, 2004.

Jorgensen, Julia, George A. Miller, and Dan Sperber. "Test of the Mention Theory of Irony." *Journal of Experimental Psychology: General* 113, no. 1 (1984): 112-120.

Kreuz, Roger J., and Sam Glucksberg. "How to be Sarcastic: The Echoic Reminder Theory of Verbal Irony." *Journal of Experimental Psychology: General* 118, no. 4 (1989): 374-386.

Kreuz, Roger J., and Richard M. Roberts. "On Satire and Parody: The Importance of Being Ironic." *Metaphor & Symbolic Activity* 8, no. 2 (1993): 97-109.

Laval, Virginie, and Alain Bert-Erboul. "French-Speaking Children's Understanding of Sarcasm: The Role of Intonation and Context." *Journal of Speech, Language, and Hearing Research* 48, no. 3 (2005): 610–620.

Rockwell, Patricia. "Lower, Slower, Louder: Vocal Cues of Sarcasm." *Journal of Psycholinguistic Research* 29, no. 5 (2000): 483–495.

Rockwell, Patricia. "Facial Expression and Sarcasm." *Perceptual and Motor Skills* 93, no.1 (2001): 47–50.

Sperber, Dan, and Deirdre Wilson. “Irony and the Use-Mention Distinction.” *Radical Pragmatics* (1981): 295–318.

Wilson, Deirdre, and Tim Wharton. “Relevance and Prosody.” *Journal of Pragmatics* 38, no. 10 (2006): 1559–1579.