THE IMPACT OF BACKCHANNELS ON THE PERCEPTION OF LISTENING COMPETENCE

A Thesis submitted to the faculty of
San Francisco State University
In partial fulfillment of
the requirements for
the Degree

Master of Arts

In

English: Linguistics

by

Jasmine Anais Rivero

San Francisco, California

May 2019

CERTIFICATION OF APPROVAL

I certify that I have read The Impact of Backchannels on the Perception of Listening Competence by Jasmine Anais Rivero, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in English: Linguistics at San Francisco State University.

Anastasia Smirnova, Ph.D.
Assistant Professor

Jenny Lederer Ph.D.
Assistant Professor

THE IMPACT OF BACKCHANNELS ON THE PERCEPTION OF LISTENING COMPETENCE

Jasmine Anais Rivero
San Francisco, California
2019

In a conversation, a listener uses backchannels to signal that he or she is attending to what the speaker is saying. Although the relationship between backchanneling and listening has been referred to substantially in the literature (Fries, 1952; Yngve, 1970; Stubbe, 1998; Ward, Escalante, Al Bayyari & Solorio, 2007; Bodie, St. Cyr, Pence, Rold & Honeycutt, 2012), this relationship has been understudied experimentally. This thesis experimentally investigates the connection between backchanneling and how a listener is perceived—specifically, if the person is perceived as a *competent listener* or not.

This thesis offers a novel methodology for experimentally researching the relationship between backchanneling and listening. The experimental section of this thesis is broken into two parts. In the first study, participants (n=70) were presented with an audio clip in which the quantity of backchannels uttered by the listener was manipulated. In the second study, an identical audio clip was presented to participants (n=30). This time the placement of backchannels was manipulated. Though the results of both experimental studies were inconclusive, the research conducted nevertheless raised

an important theoretical question, that is, how do the quantity and placement of backchannels used by a listener affect how they are perceived?

Additionally, this thesis theoretically expands on perception studies in sociolinguistics, which have traditionally focused on how the *speaker* is perceived by others (Labov, 1966; Labov, 1972; Kramer, 1977; Feldstein, Dohm, & Crown, 2001; Campbell-Kibler, 2005). In contrast, the research detailed in this thesis focuses on how the *listener* is perceived. In doing so, this thesis offers one possible way that the concept of a *competent listener* can be operationalized (based on a framework provided by Bodie et. al., 2012). In this research, A *competent listener* is someone who *understands* the speaker, *pays attention* to the speaker, and is *friendly* towards the speaker.

I certify that the Abstract is a correct representation of the content of this thesis.

Chair, Thesis Committee

05/10/2019
Date

ACKNOWLEDGEMENTS

I would like to express my gratitude to Professor Anastasia Smirnova, my research advisor, for her patience, excellent guidance, and support through the process of producing this thesis. I would also like to extend thanks to Professor Jenny Lederer for the additional support she gave me while conducting this research. I am also grateful to all my colleagues in the M.A. Linguistics program for helping me keep my head high during this process. Finally, I would like to thank my mother, Nancy Douglas, the first linguist I ever met, for teaching me the love of language and for supporting me in everything that I do.

TABLE OF CONTENTS

List of Figures viii

List of Appendices ix

1 Introduction 1

2 Background 5

2.1 Backchannels: An Introduction 5

2.2 Literature Review 14

3 Experimental Studies 24

3.1 Study 1: Manipulating Quantity of Backchannels 25

3.1.1 Method 25

3.1.2 Design and Procedure 28

3.1.3 Results 30

3.2 Study 2: Manipulating Placement of Backchannels 32

3.2.1 Method 32

3.2.2 Design and Procedure 32

3.2.3 Results 34

3.3 Discussion 36

3.3.1 Experimental Studies 36

3.3.2 Limitations 40

3.4 Summary of Findings 43

4 Methodological and Theoretical Contributions 44

4.1 Methodological Contributions 44

4.2 Theoretical Contributions 46

5 Conclusion 51

5.1 Summary 50

5.2 Directions for Future Research 51

References 52

LIST OF FIGURES

Figures Page

1. Distribution of *Uh-huh* and *Yeah*.....15
2. Distribution of Insertion Points for Backchannels..... 17
3. Mean Listening Competence Scores for Study 1.....31
4. Mean Listening Competence Scores for Study 2.....36

C.1 Training Page.....63
C.2 Introduction Page.....63
C.3 Survey Page.....64
C.4 Demographics Page.....65

LIST OF APPENDICES

Appendix Page

A. Transcripts of Recordings, Study 1 55

B. Transcripts of Recordings, Study 2 59

C. Survey Materials for Studies 1 & 2 63

Chapter 1: Introduction

Uh-huh, yeah, right. These are just some of the short utterances that are used as backchannels by interlocutors to respond to their conversation partner's discourse without interrupting said partner's flow of talk (Yngve, 1970; Ward & Tsukahara, 2000). Backchannels have long been described as a way for participants in a conversation to communicate that they are listening to their conversation partner without making moves to take over the turn (Yngve, 1970; Stubbe, 1998; Ward, Escalante, Al Bayyari & Solorio, 2007; Bodie, St. Cyr, Pence, Rold & Honeycutt, 2012). However, very little experimental research has actually been done to support this claim (i.e. that listeners use backchannels to communicate that they are listening) This paper aims to investigate the relationship between listening and backchanneling through two experimental studies in which the quantity and placement of backchannels are manipulated in a conversation. The goal is to explore how the use of backchannels affects the way in which the *listener* is perceived.

Backchannels can be verbal tokens (*uh-huh, yeah, right*) as well as non-verbal gestures (such as nodding). The research in this thesis focuses on the relationship between *verbal* backchannels and the perception of listening competence. Any reference to backchannels in this paper refers to verbal tokens only, unless otherwise noted.

This thesis also assumes that if backchannels are a way for conversation participants to signal that they are listening, they are attempting to communicate that they are listening *well*. Thus, this research examines the relationship between backchanneling

and good listening, which from this point on will be referred to as *competent listening*. Hymes (1971, p. 61) defines *competence* as "the most general term for the speaking and hearing capabilities of a person…dependent upon both (tacit) *knowledge*, and (ability for) use." Allen & Brown (1976, p. 248) emphasize that competence "is tied to actual performance of language in social situations."

The research questions motivating the experimental studies in this thesis are the following:

1. Does manipulating backchannel quantity impact the perception of how competent at listening the listener is?
2. Does manipulating backchannel placement impact the perception of how competent at listening the listener is?

These research questions are investigated through two experimental studies. The first study focuses on investigating if manipulating backchannel *quantity* impacts the perception of listener competence. The second study focuses on if manipulating backchannel *placement* impacts the perception of listener competence.

As mentioned, despite the great deal of attention that backchannels have received in the theoretical literature (Fries 1952; Yngve, 1970; Deng, 2009), there are no studies that provide empirical support for the observation that backchanneling is indicative of competent listening. This contrasts with the existence of a large body of literature that empirically investigates speaker competence and what behaviors represent competent speaking (Labov, 1966; Labov, 1972; Kramer, 1977; Feldstein, Dohm, & Crown, 2001;

Campbell-Kibler, 2005; represent just a short list of some of the literature that takes an empirical approach to investigating perceptions of the speaker).

To the best of my knowledge, there has not been any experimental research conducted that looks at the relationship between backchanneling and listening competence. Thus, I developed a novel methodology in order to determine if manipulation of the quantity and placement of backchannels has an effect on competent listening.

Bodie et. al. (2012) proposed a framework of good listening based on experimentally collected data. From this framework, I developed three questions aimed at measuring a participant's perception of a listener's competence from a short audio clip. These three questions presented to participants measured how well they thought the listener in the audio clip understood the speaker, how well the listener was paying attention to the speaker, and how friendly the listener was towards the speaker. The values collected from the three questions were averaged to determine a *listening competence score*, which was used to measure the participant's perception of the listening competence of the listener in the audio clip.

The results of the two studies were ultimately inconclusive. Due to this, I am unable to claim whether or not the quantity or placement of backchannels has an effect on the perception of a listener's listening competence. Nevertheless, the research I conducted still raises important theoretical and methodological questions related to how the quantity and placement of backchannels used by a listener impact how that listener is

perceived by others. This thesis also explores how the concept of *competent listening* should be operationalized. Finally, my research expands on perception studies of sociolinguistics that have traditionally only examined the perception of the speaker, by investigating the perception of the listener.

Overall, this research provides both a lens for examining the connection between backchanneling and listening. Additionally, this thesis provides experimentally collected data on this connection. Backchanneling and listening are important parts of conversations, and thus, human communication. Having a better understanding of the relationship between backchanneling and listening can help us strengthen our knowledge and insight into human communication.

Chapter 2: Background

2.1 Backchannels: An Introduction

Example 1 is an excerpt of a conversation between two acquaintances, *Jane* and *Mark*. Jane is telling a story about how she had to use social media to contact an acquaintance in order to send a last-minute message before an international flight. This excerpt showcases backchannels being used in natural discourse.

(1)	1 Jane:	I found him on Facebook
	2 Mark:	Uh huh
	3 Jane:	through my boyfriend's account,
	4 Mark:	Yeah
	5 Jane:	Because I had used up all the wifi at Oakland um and so I
	6	was praying
	7 Mark:	yeah yeah yeah ((nodding))
	8 Jane:	he would answer

Although the main event of the interaction in Example 1 is the story that Jane is telling (*I found him on Facebook, through my boyfriend's account, because I had used up all the wifi at Oakland um and so I, was praying, he would answer*), the other conversation participant, Mark, is not silent. During Jane's story, Mark says *uh huh, yeah, yeah yeah yeah*, as well as nods. These short utterances and one instance of gesture are examples of backchannels.

Utterances used by listeners to signal that they are paying attention to a speaker in a conversation have been referred to by different names, including *signals of attention* (Fries, 1952), *minimal responses* (Bennet & Jarvis, 1991), *minimal feedback tokens* (Gardner, 1997), *continuers* (Jurafsky, Shriberg, Fox, & Curl, 1998), and *listener responses* (Deng, 2009). This variation in terminology reflects the assumptions or elements about the functions of backchannels that the researcher wants to highlight. Most commonly, these utterances are known by the term that I will be using throughout this paper: backchannels (Yngve 1970; Ward & Tsukahara, 2000).

Backchannels share certain features with other kinds of words. Like high-effect utterances such as *ouch* and *wow* (Jackendoff, 1999), backchannels have phonological realizations and semantic meaning, though they lack robust syntax. Like discourse markers (Schiffrin, 1987), backchannels are both short in length and play an important role in the structuring of discourse. Despite these similarities, backchannels are a unique class of utterance: they have a set of distinctive features that set them apart from other types of words. These features are:

1. Backchannels are important in maintaining the flow of and creating structure in a dyadic conversation (Dittman & Llewelyn, 1967; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 1982; Kraut, Lewis & Swezy, 1982).
2. Backchannel utterances do not make attempts to take over the turn from the current speaker (Fries, 1952; Yngve, 1970).

3. Backchannels are limited exclusively to interactive speech between at least two or more interlocutors.
4. Backchannels have a specific and characteristic intonation contour (Jurafsky et. al., 1998; Beňuš, Gravano, & Hirschberg, 2007).

The first distinguishing feature of backchannels is that they are important in maintaining and creating structure in dyadic conversation. They are a part of a listener's active participation in the conversation (Schegloff, 1982) and help to ensure that an exchange goes smoothly (Dittman & Llewelyn, 1967). Backchannels allow listeners to signal to the speaker that they are actively participating (i.e., paying attention) and are interested in continuing and maintaining a conversation (Kraut et. al., 1982) They signal that the speaker should continue talking. At given points in a conversation, listeners can also use backchannels to pass on an opportunity to take over the turn (Sacks et. al., 1974). In this way, backchannels are important for creating structure in a conversation.

The second feature of backchannels is that they do not indicate that the listener is making any attempts to take over the turn from the current speaker. In Example 1, Mark's *uh-huh* and *yeah* do not interrupt or halt Jane's narrative from continuing. Jane continues her story, her turn, right through Mark's backchannels. There is no expectation from either interlocutor that there is going to be a change in turn, as can be seen through Jane's continuous storytelling.

The third feature of backchannels is that they are limited exclusively to synchronous interactive speech between at least two or more interlocutors. Backchannels

are inherently a feature of simultaneous and spontaneous interaction. The speech of a person who is alone, talking aloud to themselves, will not contain backchannels. Once a second interlocutor is added to this interaction, backchanneling is used by the other person in order to participate in the conversation.

The final characterizing feature of backchannels is that they have a unique prosody, which can be used by speakers as a way to distinguish them from other types of utterances like agreements or discourse markers, especially if a specific lexical token can be realized as a backchannel or a different type of utterance (Jurafsky et. al., 1998; Beňuš et. al., 2007). Specifically, backchannels have a higher pitch, intensity, and pitch slope than other types of utterances (Beňuš et. al., 2007, p. 1067).

Methods of approaching the study of backchannels have been varied; however, literature on backchannels can be split into two approaches, referred to by Deng (2009) as the *lumping* approach, which is most often used in the fields of linguistics, communication studies, and psychology, and the *splitting* approach, which is primarily used in research done in the field of conversational analysis. The *lumping* approach examines the functions, roles, and structural properties of backchannels examined as a single unified class of utterances. The splitting approach examines the functions of specific backchannels as they appear in context as used in spoken language, usually with an emphasis on the role that individual backchannels play in relation to turn-taking, or how interlocutors speak to each other in alternating turns (Sacks et. al., 1974). This section will focus primarily on literature that falls under the *lumping* approach.

Fries (1952) was one of the first scholars to describe backchannels and refers to them as *signals of attention*, classifying them as a unique type of utterance. According to Fries, these signals of continued attention are brief vocalizations that do not interrupt the talk of the main speaker. Most significantly, Fries points out that signals of continued attention "signal that [the listener] is listening attentively to the speaker" (p. 50). This early description highlights that backchannels are used for a person to demonstrate to their conversation partner that they are listening.

The term *backchannel* was first used by Yngve (1970) as the *back channel*. Yngve used the word to distinguish the *back* channel of a conversation from the *main* channel of the conversation, where one of the interlocutors in the dyad has the floor and is speaking in turn. In Example 1, Jane's speech would be described as the *main channel* while Mark's *uh-huh* and *yeah* are a part of the *back* channel. The goal of this paper was to describe the functions of this type of utterance as a whole.

Yngve (1970) described the backchannel as where "the person who has the turn receives short messages such as *yes* and *uh-huh* without relinquishing the turn. The partner, of course, is not only listening but speaking occasionally as he sends the short messages in the back channel" (p. 568). Yngve's description of there being a "back" channel that that exists concurrently with the "main" channel of the conversation is inherently describing a collaborative form of dialogue. Conversations are more complex than just one person speaking while the other person is listening silently; instead, the listener is giving the speaker verbal and gestural feedback throughout their turn, allowing

for a steady flow of communication. A part of the description of backchannels from Fries (1952) was that they are utterances that do not interrupt the flow of the speaker. Yngve's description of backchannels, adds to this characteristic of backchannels: they do not interrupt the other speaker's *turn*.

Backchannels have been described as being a useful tool for speakers to gauge whether their conversation partner is listening to them. Yngve describes two scenarios that can happen between two people if the listener is not using backchannels. The first is of two people talking on the telephone. If there is a lack of backchannels from the listener, the other person will "come to a grinding halt and say something like, 'Hello, are you still there?'" (p. 568). In the second scenario, a husband, although listening to what his wife is saying does utter any backchannels and thus is "accused of not paying attention to what his wife is saying although he has heard and understood her every word" (p. 568). By exploring two scenarios of what can happen when backchannels are not used, we can see how backchannels have been described to play an important role in indicating listening and attention to the person speaking.

Besides being used by speakers (or the non-backchanneling person) as a way to monitor the listening and attention of the person they are speaking to, backchannels have been found to have other effects on speakers. When there is a delay in audible backchannels, speakers have more difficulty encoding their information (they use a greater number of words in their encoding process) (Krauss & Weinheimer, 1966). When a speaker is narrating a story, lower numbers of *specific responses* (backchannels that

relate to specific content in the narrative) from the listener led to the quality of the narrator's stories suffering (Bavelas, Coates, & Johnson, 2000). Backchannels can have an even more overt effect on a speaker's narrative than just impacting the quality of a narrator's story. Different types of backchannels change the development of the narrative of the speaker in different ways. Backchannels that responded in some way to the content of the narrative such as *oh* and *wow* lead a speaker to elaborate on the previous event mentioned while backchannels that are more generic like *mhm* and *uh-huh* lead to the narrator to move on to a new event in the story (Tolins & Fox Tree, 2014).

Within contemporary literature on backchannels that focuses on the listener (the person backchanneling), there seems to be a consensus that backchannels are representative of good listening (Ward et. al., 2007; Poppe, Truong, & Heylen, 2011; Bodie et. al. 2012), which is in line with early descriptions of backchannels. There is not a common way of operationalizing what *good listening* is across this literature. Ward et. al. (2007) created a method for second-language learners to acquire native-like backchanneling behavior in Arabic with the motivation that "to be a good listener you have to be able to show you are listening," and that "good listeners generally produce back-channel feedback." In other words, *showing* that you are listening is a feature of good listening, and you can show that you are listening through backchanneling.

Ward et. al.'s (2007) claim evokes the scenario described from Yngve (1970): even though the listeners he described (the person on the phone and the husband) were listening, they weren't performing *competent listening* because they were not producing

backchannels. Those lack of backchannels led the speakers in Yngve's scenarios to halt their speech in order to explicitly check that their conversation partners were listening. Poppe et. al. (2011) chooses to focus on *understanding* as the feature of good listening to investigate. Poppe et. al. describe that listeners "actively [contribute] to the conversation by signaling attention, interest and understanding to the speaker" by using backchannels. In a preliminary framework of an Implicit Theory of Listening that enumerates qualities and behaviors that characterize *good listening*, backchanneling is one of those behaviors (Bodie et. al., 2012). As can be seen by how different each of these three studies chooses to characterize *good* or *competent* listening, there is a lack of consistency in how *competent listening* is operationalized in the past literature.

Despite the numerous descriptive accounts in both early and more recent literature that backchannel feedback is an indication of not only listening but *competent listening* behavior, there is a gap in the literature of experimental data that supports this claim. In an experimental investigation of the communicative function of backchannels, it was found that backchannels are used to "denote agreement" and to "suggest a context of informality" (Bennett & Jarvis, 1991), with no significant findings related to listening. By examining the connections between the perception of listener competence and backchannels in this paper, I will be contributing concrete data regarding the relationship between backchannels and listening.

This section provided an introduction to the topic of backchannels through a review of the literature on the topic that is considered to be taking the *lumping* approach

to the study of the phenomenon: literature concerned with the functions, roles, and structural properties of backchannels as a class of utterances. Backchannels are short utterances (tokens such as *uh-huh* and *yeah*) spoken by a listener that does not interrupt the turn of the main speaker. Backchannels play an important role in discourse, impacting the quality, quantity, and type of content the speaker utters. Finally, despite the lack of experimental data to support the connection, there is a significant amount of descriptive literature that backchannels are an important part of good listening, and are used by both the backchanneler to communicate their listening and attention to the main speaker and by the main speaker to gauge the quality of listening their conversation partner is providing them in their interaction.

2.2 Literature Review

This section presents literature that is specifically relevant to my experimental studies. Firstly, this section provides data for the most common lexical realizations of backchannels in American English. Then, this section reviews literature that provides data on where backchannels occur in relation to the utterances of the main speaker. Thirdly, this section highlights research that provides data in relation to the manipulation of quantity of backchannels. Finally, this section looks at literature that provides a framework of what *competent listening* is, as well as if and how backchanneling is connected to *competent listening*.

Backchannels in American English are most commonly realized through a specific set of lexical items. Jurafsky et. al. (1998) provide data on the lexical realizations

of backchannels, as well as specifying the differences between different types of backchannels. To gather data on the features of backchannels, Jurafsky et. al. manually tagged and analyzed the Switchboard-Discourse Annotation & Markup System of Labelling Corpus (SWBD-DAMSL) (Jurafsky, et. al. 1997, as cited in Jurafsky et. al., 1998). The researchers used 42 tags in total; however, the most relevant tags for our discussion are the four related to backchannels: *continuer, agreement, incipient speaker,* and, *yes-answer*. Jurafsky et. al. consider these four tags to all be subsumed under the umbrella term *backchannels* or *acknowledgment tokens. Continuer*-backchannels, or utterances that are "indicating that the other speaker should go on talking" (p. 115), are the most common type of backchannel, as well as the type of backchannel thtis research is most concerned with.

The four types of backchannels have significant overlap in lexical tokens with words like *yeah, uh-huh, yes,* and *right,* being some common examples. Although there are words in common across the four backchannel types identified by Jurafsky et. al. (1998), each type has a particular distribution of use of different tokens. Figure 1, reproduced from Jurafsky et. al. provides a summary of the distribution of the tokens *uh-huh* and *yeah* for each of the four backchannel types.

Agreements		**Continuer**		**Incipient Speaker**		**Yes-Answer**	
yeah	36 %	*uh-huh*	45%	*yeah*	59%	*yeah*	56%
uh-huh	4%	*yeah*	27%	*uh-huh*	17%	*uh-huh*	14%
other tokens	60%	other tokens	28%	other tokens	24%	other tokens	30%

Figure 1: Distribution of Uh-huh *and* Yeah

Uh-huh is the most frequent lexical token for *continuer*-type backchannels, the type of backchannels of most interest to this research. *Uh-huh* is used as a *continuer*-type backchannel approximately 2.5 times more often than used by *incipient speaker* type backchannels, approximately 3 times more often than by *yes-answer* type backchannels, and approximately 11 times more often than by *agreement* type backchannels.

Although early descriptions of backchannels, like from that of Fries (1952), describe the location of backchannels as unpredictable, work from Dittman & Llewelyn (1967), Ward & Tsukahara (2000), Cathcart, Carletta & Klein (2003), and Kjellmer (2009) provide both descriptions and corpus data that demonstrate the systematicity of backchannel occurrence in English.

Dittman & Llewellyn (1967) demonstrate that backchannels *do* indeed have systematicity in where they occur. Dittman & Llewellyn used *listener responses* (their term of choice for referring to backchannels) as a way to provide evidence that spoken language is decoded by the listener in phonemic clauses. In their research, Dittman & Llewellyn found that listener responses tended to follow phonemic clause endings, and that listener responses rarely occur phonemic clause medially. Although this research is not directly focused on backchanneling, the paper offers important experimental data and insight into the occurrence of backchannel utterances in relation to a main speaker's speech.

Ward and Tsukahara (2000) provide extensive evidence that shows that regions of low pitch late in an utterance seem to act as a cue for backchannel feedback. Ward &

Tsukahara (2000) found that "low-pitch regions lasting 110 ms" in both English and Japanese serve as a cue for backchannels. Low-pitch regions of 110 ms frequently accompany grammatical completion. Ends of utterances and clauses tend to be marked with low pitch, thus the findings of this research support that backchannel feedback comes after clause and utterance endings.

Cathcart et. al (2003) used the HCRC Map Task Corpus of Scottish English to create models to predict backchannel location. In short, transition relevance points (Sacks et. al., 1974) are locations where hearers have the opportunity to take up the main channel of discourse if they are interested in doing so. If they do not, and instead give feedback by using a backchannel, and the current speaker must decide if they want to continue their turn. Thus, turn relevance locations can be considered possible backchannel insert points.

The finding that backchannels occur at the ends of turns is reproduced in the corpus work done by Kjellmer (2009). Kjellmer used a corpus-based approach to answer the question, *where do we backchannel?* Kjellmer's study uses material from the CobuildDirect Corpus of spoken British English. Although it is common for backchannels to come directly after an intonation unit, backchannels also do occur before the end of a speaker's turn. The majority of backchannels are what Kjellmer refers to as *turn external* backchannels, these are backchannels that occur after the end of the speaker's turn. The end of turns tends to correspond with the end of clauses and utterances. Backchannels that are not *turn external* backchannels are referred to as *turn*

internal backchannels, or backchannels that occur during the turn of the main speaker. The majority of turn internal backchannels, 45.4%, occur at the end of a clause break.

Is there a significant difference between *turn external* backchannels and *turn internal* backchannels that occur at clause breaks? Turns are completely arbitrary in length and can be a word long or hundreds of words long. Both these types of backchannels defined by Kjellmer (2009) can occur at clause breaks, which supports the findings from Ward & Tsukahara (2000) that backchannels are frequently found at points of low pitch that frequently occur clause and utterance finally. Regardless if Kjellmer's decision to make a distinction between these two types of backchannels is warranted, his data shows that clause-final backchannels occur most commonly in the corpus. Figure 2 reproduces some data from Kjellmer (2009).

Insertion Point	Percentage
At clause break	45.4%
In a prepositional phrase	13.2%
In a coordinated structure	8.8%
In a verb phrase	8.2%
In a noun phrase	5.4%
Adverbial	5%
Before a repeated element	4.5%
In nexus	4.1%
Unanalyzed	2.9%
Before apposition	2.1%
In an adjective phrase	0.4%
Total	**100%**

Figure 2: Distribution of Insertion Points for Backchannels (Kjellmer, 2009)

People are sensitive to the quantity of backchannels used during an interaction. Higher frequency of backchannels has been found to lower the enjoyment of a conversation (Li, Cui, Wang, Leske, & Aguilera, 2010), while too many or too few

backchannels per minute have been found to impact how human-like a listener is perceived (Poppe, Truong & Helyen, 2011).

Li et. al. (2010) set out to investigate the differences of backchanneling behavior between two different groups: native English-speaking Caucasian-Canadians, and English proficient persons originally from mainland China who had been in Canada for an average of 27 months. The only significant finding from the study in regards to participant enjoyment of the conversation was that there was a negative correlation between the frequency of backchannel responses and participants level of enjoyment of the conversation. Li et. al. claim that this data suggests that the more frequent backchannel responses are, the lower the enjoyment of the conversation is. Specifically, their study investigated speaker presentation, listener recall, and participant enjoyment of the conversation in an intercultural setting. 47 dyads (7 eliminated due to experimental error) were split into conditions made up of same-gender intercultural pairs prompted to simulate a physician-patient interview. Each interaction was split into two sections: the first section involved the patient providing information about their case history to the physician while the second section involved the physician instructing the patient on how to use a drug. Immediately after the dialogues had concluded, participants were given questionnaires to fill out asking about their experience of the interaction.

The implications of this finding are not particularly straightforward due to the vagueness of the variable "enjoyment of the conversation." Li et. al. (2010), do not provide a definition for this parameter in their investigation; however, we can speculate

that *enjoyment of the conversation* might refer to the positive emotion or impact that occurs in reaction to a sequence effective communication. If *enjoyment* is related to the effectiveness of the communication between a dyad, then the high frequency of backchannels might be associated with ineffective communication.

Research from Poppe et. al. (2011) supports the findings from Li et. al (2010) that too many backchannels are associated with negative perception as well as adding that too few backchannels also impacts the perception of the quality of a listener. The goal of the investigation conducted by Poppe et. al was to gather information regarding optimal quantity, timing, and type of backchannels produced by a listener in order to effectively develop an AI listener. In their experiment, n=24 participants viewed stimuli of a human speaker side-by-side with an animated listener and were asked to evaluate how human-like they perceived the behavior of the artificial listener. The backchanneling behavior of the artificial listener was systematically manipulated across the following dimensions: quantity (the number of backchannels uttered), type (either the backchannels were animated as were produced by the human listener or the type, verbal, non-verbal, or both, were switched), and timing (timings as were in the original human recording and random timings). By crossing the three dimensions, 12 experimental conditions were yielded.

Participants evaluated how human-like they perceived entire dialogue fragments as well as individual backchannels. An analysis of both produced the following conclusions: for timing, random timing was perceived as less human-like than the original timings; for type, switched timing was perceived as less human-like than the

original types. The most interesting findings from Poppe et. al. were from the dimension of quantity. In conditions that used the original (human) timings and types backchannels, more backchannels led to the fragment being perceived as more human-like; however, *too* many backchannels and the ratings of the fragment lowered. The researchers conclude from these observations that there is both a lower and an upper bound on the number of optimal backchannels per minute, they say between 6 and 12 backchannels per minute. The findings that there is a lower and upper bound for the number of optimal backchannels, and that these bounds are 6 to 12 backchannels per minute is an important concrete qualitative contribution to the understanding of backchannels.

The quality "human-like," used by Poppe et. al. (2011) to describe the quality of artificial listeners is not specifically elaborated on or described in their research. In order to be able to make the findings from Poppe et. al. relevant outside just their research question, it is necessary to identify what "human-like" means. If the goal of the researchers was to create a listener that is as *human* as possible in order to optimize the communication between a human and an AI-agent, it is safe to assume that the human behavior the AI is modeled after would be an example of a highly competent communicator, and in this case, a highly competent listener. Thus, we can conclude that the most "human-like" listener would be the most competent listener. I hypothesize that the most "human-like" listener would also be the most "competent listener."

What does *competent listening* look like? Overall, there is a lack of experimentally collected data and literature on the topic of competent listening; however,

a preliminary framework of an Implicit Theory of Listening based on experimentally collected data was proposed by Bodie et. al. (2012). This framework outlines the characteristics or attributes of competent listening, as well as the behaviors that listeners do to demonstrate the characteristics of competent listening. Backchannel responses are one of the behaviors that listeners use in order to do competent listening.

In order to create their Implicit Theory of Listening framework, Bodie et. al. (2012) conducted three studies. The goal of the first study was to generate a list of both general and specific behaviors from participants that they easily associate with competent communication and to figure out which (if any) of these behaviors are related to competent listening. The second study was conducted in order to cross-validate the relationships created by researchers between each response and listening competence and each category and listening competence. The final study aimed to do two things: first, to replicate the meaning ratings of the listening competence for each identified behavior and second, to determine which behaviors participants considered most related to their impressions about listening.

The methods of the first study involved using a computer, participants were provided with text boxes where they were asked to input one characteristic or behavior per box that they felt contributed to concluding that someone is communicatively competent. After listing the behaviors, the computer generated the responses they had given and then asked participants to rate the responses on a 6-point scale with (6) being 'definitely a characteristic of competent listening' and (1) being 'definitely *not* a

characteristic of listening competence.' In the second study, participants were assigned randomly to one of 6 conditions where they were asked to list behaviors that would have lead them to think that someone was: understanding; a good listener, paying attention; enabling conversational flow, or friendly. After listing the behaviors, the computer generated the responses they had given and then asked participants to rate the responses on a 6-point scale with (6) being 'definitely a characteristic of competent listening' and (1) being 'definitely *not* a characteristic of listening competence.' After rating the behaviors, participants were directed to an online survey where they were asked to complete a scale asking them to assess how related the categories coming from Study 1 were related to the impression of the other person as a listener. In the final study, participants were given three separate questionnaires.

From their data collection, Bodie et. al. (2012) developed a framework for an Implicit Theory of Listening. Implicit theories are "implicit theories of listening are proposed as knowledge structures that specify sets of interconnected listening-related attributes and are used during interactions, especially first encounters, to infer an individual's competence in listening from observed behavior" (p. 21). The framework is made up of five attributes and nineteen behaviors connected to those attributes that make up the schema of a *good listener*. The five attributes that are a part of the framework are *pays attention*, *friendly*, *responsive*, *conversational flow*, and *understanding*. Backchannels are one of the nineteen behaviors that make up the *good listener* schema. The researcher's definition of backchanneling is very much in-line with the definitions of

backchannels that have been brought up so far in this literature review: "saying uh-huh and yeah to signal they understand you" (p. 17). In their data analysis, they found as a behavior, backchannel responses are most associated with the attributes of responsiveness and conversational flow. Although backchannel responses are *most* associated with these two attributes, it does not mean that they are not necessarily associated with the attributes of paying attention, friendliness, or understanding.

Chapter 3: Experimental Studies

In an effort to gather data about the connection between backchanneling and listening competence, I created two studies, one manipulating the quantity of backchannels uttered, and another manipulating the placement of where the backchannels were uttered (in relation to the speech of the primary speaker). The questions that these studies aimed to answer were:

1. Does manipulating backchannel quantity impact the perception of how competent at listening the listener is?

2. Does manipulating of backchannel placement impact the perception of how competent at listening the listener is?

For the first study, I hypothesized that manipulating the quantity of backchannels would impact the perception of the listening competence of the listener. Specifically, I expected a listener uttering quantities of backchannels that fall within the ideal range identified by Poppe et. al. (2011) to be perceived as the most competent. Given the findings from Poppe et. al. and Li et. al. (2010), that there is an upper limit for an ideal number of backchannels and that number of backchannels is correlated with a lack of enjoyment of a conversation, respectively, I hypothesized that a listener uttering quantities of backchannels that fall above the ideal range would be considered the *least* competent. A listener uttering quantities of backchannels that fall below the ideal range would be considered more competent than the listener uttering quantities above the ideal

range, but less competent than the listener uttering quantities of backchannels within the ideal range.

For the second study, I hypothesized that manipulating the placement of backchannels would impact the perception of the listening competence of the listener. Specifically, I expected a listener backchanneling at a clause break to be the most competent listener. This hypothesis is based off various research that found that backchannels most frequently occur at clause breaks (Dittman & Llewelyn, 1967; Ward & Tsukahara, 2000; Cathcart et. al., 2003; Kjellmer, 2009). I expected that a listener who backchannels inside of a noun phrase to be perceived as a less competent listener than the listener who backchannels at clause breaks. This was motivated by the fact that backchanneling inside of noun phrases is found to occur, according to Kjellmer (2009), but at a rate much lower than backchannels that appear at clause breaks. Finally, I expected a listener that backchannels word medially (in the middle of a word spoken by their conversation partner) to be perceived as the least competent at listening. This was motivated by the fact that backchanneling word medially is a location that has not been described as a backchannel insertion point in English.

3.1 Study 1: Manipulating Quantity of Backchannels

3.1.1 Method

Participants. Seventy participants were recruited through the crowdsourcing marketplace Amazon Mechanical Turk to participate in Experiment 1. Amongst the participants, 58.62% were men and 41.38% were women. The average age of participants

was 36.53 years (SD = 10.66). Three additional participants were excluded from the data set for not speaking American English as their dominant language, three participants were excluded from the data set for not completing the survey, and another nine additional participants were excluded from the data set for completing the survey without listening to the experimental stimuli. It was determined that a participant had completed the survey without listening to the experimental stimuli by looking at the amount of time the participant took to complete the survey. When added together, the audio tracks that participants needed to listen to totaled 210 seconds in length. If a participant had completed the survey (had answered all the questions) in less than 210 seconds, they were excluded, as it would have been impossible for them to have listened to the experimental stimuli audio tracks in full.[1]

---

[1] Data was collected in two separate stages. I first collected data from 30 participants. However, the sample was unbalanced with 72% males and 28% females. After an analysis of the data from this first collection stage, it was found that there was a significant difference between the 9 backchannels condition and the other conditions combined between males and females after conducting an unpaired t-test $t(18) = 2.40$, $p = 0.03$. This difference could have been attributed to the imbalance of males and females in the sample. In order to determine if there was a significant difference in the responses between males and females and that the difference was not just a result of a Type I error, more data was collected in a second stage. When the data from the two stages were considered together and an unpaired t-test was conducted between the differences of the 9 backchannels condition and the other conditions between Males and Females, it was determined that the original p value was a Type I error $t(43) = 0.74$, $p = 0.46$.

Materials. Study 1 was conducted using a survey to present the experimental stimuli and questions to participants. The surveys were created using the online software Qualtrics and distributed to participants using the crowdsourcing platform Amazon Mechanical Turk (AMT).

The recordings used as the experimental stimuli were recorded using a PC laptop computer and edited using the open-source editing software Audacity.

Experimental Stimuli (Audio). The experimental stimuli are framed as a conversation between two people. In the audio, two classmates speak to each other. One, will be referred to here on out as the *narrator*, named *Lisa* in the dialogue, is telling a narrative to the other classmate, referred to here on out as the *listener*, is named *Jen* in the dialogue.

Each audio track was approximately 1 minute and 10 seconds in length. The beginning of each audio track consisted of a quick orienting dialogue between the two speakers in order to put the experimental portion of the dialogue in context (See appendix A). The experimental stimuli consisted of the same audio track for each condition. (see appendix A.) The experimental portion of the dialogue (the portion of the audio track not including the short orientation) was approximately 65 seconds in length.

3.1.2. Design and Procedure

Independent Variables. In Study 1, there was one independent variable: number of backchannels. Three conditions were developed, a 3-backchannel condition where the listener uttered 3 backchannels total, a 9-backchannel condition, where the listener

uttered 9 backchannels total, and a 15-backchannel condition, where the listener uttered 15 backchannels total. The number of backchannels for each condition was based off the findings from Li et. al. (2010) that too many backchannels is perceived as unpleasant for conversational participants and the findings from Poppe et. al. (2011) that there is a lower and upper bound for an acceptable number of backchannels per minute (6 and 12, respectively). The three independent variables were turned into the following conditions:

3-backchannel condition. This condition was designed to have a number of backchannels that is *less* than Poppe et. al.'s proposed range.

9-backchannel condition. This condition was designed to have a number of backchannels that is *within* Poppe et. al.'s proposed range.

15-backchannel condition. This condition was designed to have a number of backchannels that is *above* Poppe et. al.'s proposed range.

Dependent Variables. The dependent variables for this experiment were designed with the goal of measuring the participant's perception of the *listening competence* of the listener during the dialogue. Three variables, *understanding*, *(pays) attention* and *friendliness* were chosen based off of the preliminary framework proposed by Bodie et. al. (2012) for an implicit theory of listening in order to measure listening competence.

The dependent variables were framed to participants using the following questions: *How well did the listener (Jen) understand the narrator's (Lisa) story? How well was the listener (Jen) paying attention during the narrator's (Lisa) story?* and *How friendly was the listener (Jen) during the narrator's (Lisa) story?* (See appendix C).

Control Variables. In order to prevent incorrect data results, a variety of factors were controlled for across conditions.

Backchannel Token. The backchannel used in the stimuli was *uh-huh.* This decision was based on the finding presented by Jurafsky et. al. (1998) that the token *uh-huh* is the most common lexical realization for continuer backchannels.

Distribution of backchannels. Backchannels were distributed as evenly as possible throughout the dialogue for each condition (see appendix A for scripts of each condition to see where backchannels were inserted.)

Other variations. In order to control for any possible variations in the backchannels such as pitch, intonation, or length, the same clip of the person acting as the listener saying *uh-huh* was copied and edited into the audio track of the main dialogue of the speaker.

Procedure. After providing informed consent, participants were presented with a question that had the same format as the experimental questions accompanied by an example audio track to help familiarize them with the layout of the questions to follow. The audio was a 4-second dialogue performed by the same speakers who perform the audio that serves as the experimental stimuli. The practice question (see appendix C) for the participants was a qualitative question that was similar in style to the experimental questions. Participants were prompted to respond using a slider marked 0-100 with 0 being *Not Polite* and 100 being *Very Polite*. All subsequent experimental questions used the same format to present the experimental stimuli and questions to participants. Prior to

the experimental section of the survey, participants were briefed with a paragraph describing the task that they would be expected to complete (see appendix C).

Each participant was presented with each of the three conditions in randomized order. After listening to the experimental stimuli, participants were prompted with three questions (listed in the Dependent Variables section) aimed at determining how they rated the listening competence of the listener in the dialogue.

3.1.3 Results

In Study 1, three numerical values were collected for each condition on a scale from 0-100. These values were for the variables of *understanding*, *attention,* and *friendliness*. By finding the average value for these three variables, a value for *competent listening* was computed for each condition per participant. Since it was hypothesized that listeners who used 9 backchannels a would be perceived as being more competent listeners than those using 3 or 15 backchannels, the values for the 3- backchannel and 15-backchannel conditions were collapsed for the purposes of analysis.

A paired t-test was conducted to compare the perception of listening competence in the conditions of a listener in the 9-backchannel condition ($M = 60.39$, $SD = 20.48$) and the other two conditions combined ($M = 60.43$, $SD = 17.44$). Results were not significant between the two groups $t(57) = -0.02$, $p = 0.99$; $d < 0.01$. The effect of this analysis ($d < 0.01$) was found to be much smaller than Cohen's (1988) convention for a small effect.

When comparing the conditions of 3 (M = 60.5, SD = 19.26) and 9 backchannels using a paired t-test, the results were non-significant $t(57) = 0.04$, $p = 0.97$. When comparing the conditions of 3 and 15 backchannels (M = 60.35, SD = 21.27) backchannels using a paired t-test the results were non-significant $t(57) = .05$, $p = 0.96$.

Finally, when comparing the conditions of 9 and 15 backchannels, the results were non-significant in favor of 9 backchannels $t(57) = 0.01$, p =0.99. Figure 3 provides a visualization of the means of the 3-, 9-, and 15-backchannel conditions compared.

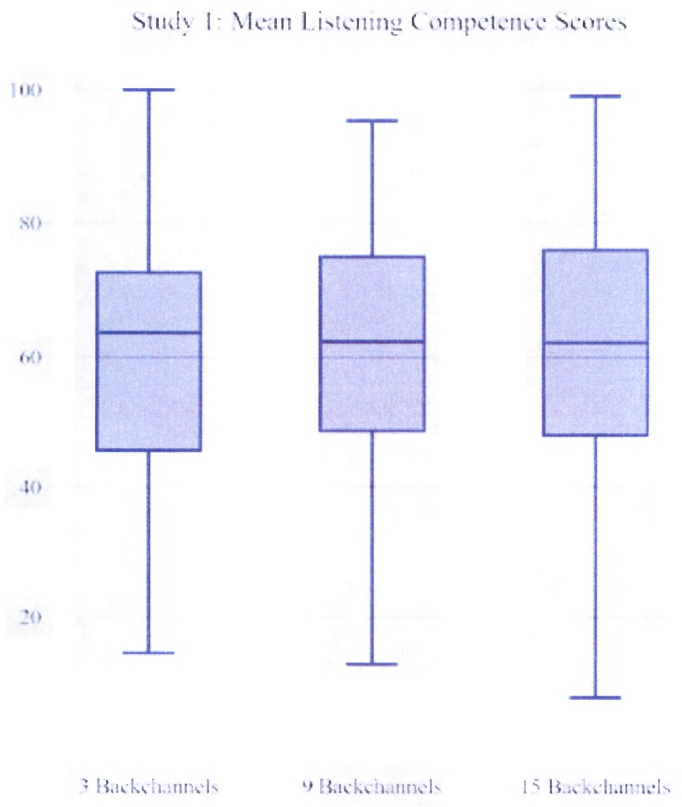

Figure 3: Mean Listening Competence Scores for Study 1

3.2 Study 2: Manipulating Placement of Backchannels

3.2.1 Methods

Participants. Thirty participants were recruited through the crowdsourcing marketplace Amazon Mechanical Turk to participate in Study 2. Amongst the participants, 76.67% were men and 23.33% were women. The average age of participants was 36.5 years (*SD*=10.71).

Materials. The materials used for Study 2 were the same as the materials used in Study 1. A survey was developed using Qualtrics and distributed through AMT. The base audio for the experimental stimuli (a conversation between speaker *Lisa* and listener *Jen*) used in this study is the same as that in Study 1, with a difference in the independent variables (backchannels).

3.2.2 Design and Procedure

Independent Variables. The independent variable manipulated in Study 2 was the placement of backchannels in relation to the speech of the speaker. The goal in choosing the independent variables was to develop three conditions wherein one reflected the most common location that backchannels occur in English, one that reflected a location where backchannels can occur in English, but less frequently, and finally one that reflected a location where backchannels have not been described to occur in English. The three variables chosen were turned into the following conditions.

Clause break condition: This condition was designed to reflect the most common backchannel insertion point in English (Dittman & Llewelyn, 1967; Ward & Tsukahara, 2000; Cathcart et. al., 2003; Kjellmer, 2009).

Noun phrase condition: This condition was designed to reflect a possible, but less occurring backchannel insertion point in English (Kjellmer, 2009).

Word-medial condition: This condition was designed to reflect a backchannel insertion point not possible in English, based on information available about insertion points.

Dependent Variables. The dependent variables used in Study 2 are identical to the dependent variables used in the first study. For ease of reference, the questions created based on the dependent variables are the following: *How well did the listener (Jen) understand the narrator's (Lisa) story? How well was the listener (Jen) paying attention during the narrator's (Lisa) story?* and *How friendly was the listener (Jen) during the narrator's (Lisa) story?*

Control Variables. In order to prevent incorrect data results, a variety of factors were controlled for across conditions. The choice of backchannel token for Study 2 is the same as the one used in Study 1, *uh-huh* (Jurafsky et. al., 1998). Like in Study 1, backchannels were distributed as evenly as possible throughout the dialogue for each condition (see appendix B for scripts of each condition). Like in Study 1, any other variations in the backchannels were controlled by using the same clip of the person acting

as the listener saying *uh-huh* was copied and edited into the audio track into the main dialogue of the speaker.

Backchannel number. The number of backchannels was controlled for across conditions in Study 2. The choice to use nine backchannels for each of the conditions stems from the Poppe et. al. (2011) paper that served as a jumping off point to determine the values in Study 1. As discussed earlier, Poppe et. al. find that the lower and upper limits of backchannels per minute are 6 and 12 respectively. By taking the average of the lower and upper limits presented by Poppe et. al., I derived the number 9. Since Study 2 is manipulating the location of the backchannels, it having each condition have the consistent number of 9 backchannels total, a number that is within the bounds of an appropriate number of backchannels ideally reduces the number of confounding factors.

Procedure. The procedure for Study 2 closely resembled the procedure used in the first study. After providing their informed consent, participants were provided with a training question. The training question presented to participants in Study 2 was identical to that of the first experiment. Prior to the experimental section of the survey, participants were briefed with a paragraph describing the task that they would be expected to complete (see appendix C).

Each participant was presented with each of the three conditions in randomized order. After listening to the experimental stimuli, participants were prompted with three questions (listed in the Dependent Variables section) aimed at determining how they rated the listening competence of the listener in the dialogue.

3.2.3 Results

In Experiment 2, I collected 3 numerical values for each condition on a scale from 0-100. These values were for the variables of *understanding*, *attention,* and *friendliness*. By finding the average value for these three variables, a value for *competent listening* was computed for each condition per participant. Since it was hypothesized that listeners who used backchannels at clause breaks would be perceived as being more competent listeners than those who used backchannels at noun phrases or word medially, the values for the noun phrase and word-medial conditions were collapsed.

A paired t-test was conducted to compare the perception of listening competence in the conditions of a listener backchanneling at clause breaks to a listener backchanneling at other locations (inside noun phrases and word medially). Results indicated a non-significant trending in the predicted direction $t(29) = 0.65$, $p = 0.52$; $d = 0.05$, indicating that a listener backchanneling at clause breaks ($M = 53.09$, $SD = 24.32$) was perceived to be a better listener than listeners backchanneling in other locations ($M = 51.61$, $SD = 25.06$). The effect of this analysis ($d = 0.05$) was found to be much smaller than Cohen's (1988) convention for a small effect.

A comparison of backchanneling at clause breaks to backchanneling within noun phrases ($M = 50.86$, $SD = 24.32$) and of backchanneling at clause breaks to backchanneling in the middle of words ($M = 52.43$, $SD = 28.11$) returned non-significant results. $t(29) = 0.91$, $p = 0.37$ and $t(29) = 0.23$, $p = 0.81$.

When comparing backchanneling within noun phrases to backchanneling in the middle of words, the results were non-significant. $t(29) = 0.55, p = 0.81$. Figure 4 provides a visualization comparing the means of the three individual conditions.

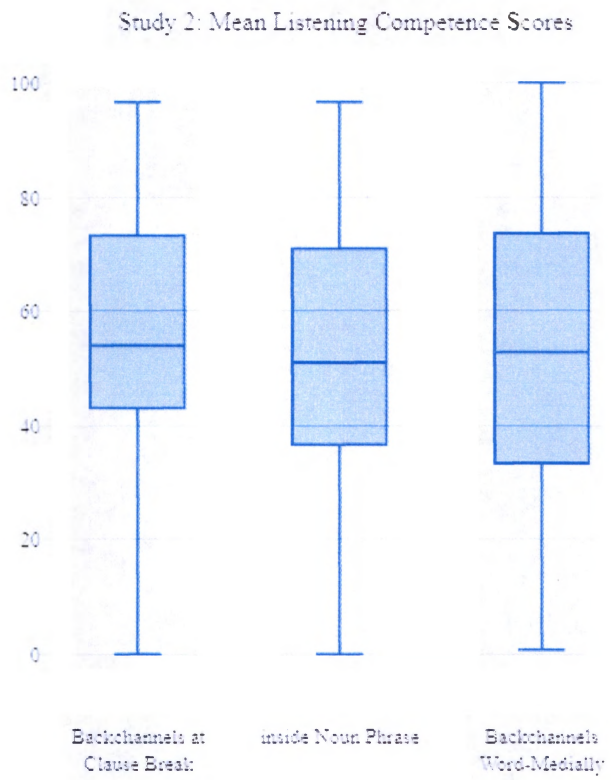

Figure 4: Mean Listening Competence Scores for Study 2

3.3 Discussion

3.3.1 Experimental Studies

For Study 1, it was hypothesized that manipulating the number of backchannels would have an effect on participants' perception of the listening competence of the backchanneler. Basing off research from Poppe et. al. (2011) and Li et. al. (2010) it was hypothesized that the 9-backchannel condition would be given the highest rating for

listening competence, followed by the 3-backchannel condition, with the 15-backchannel condition being rated the lowest for listening competence. My original hypothesis was based in literature from Poppe et. al. and Li et. al., both claimed that there is a range for the ideal number of backchannels appropriate for a listener to utter. Poppe et. al. claimed that this range would be between 6 and 12 backchannels per minute. While Li et. al. did not offer specific numbers, they, like Poppe et. al., they stressed that the use of both too *many* as well as too *few* backchannels has a negative impact on the interaction. Ultimately, no results from Study 1 were significant. There are some possibilities as to why no results from Study 1 were not significant. Firstly, manipulation of backchannel quantity might not have a meaningful effect on perceptions of the listener in intra-cultural communication (Poppe et. al; Li 2006) Second, the research questions and goals of Study 1 and the research that the experiment was designed around may have been too different, making applying the findings from Poppe et. al. not useful.

Manipulation of backchannel quantity might not have a meaningful effect on the perception of the listener, given the fact that the conditions from Study 1 all returned non-significant results. Poppe et. al (2011) found that the number of backchannels used per minute was not a significant factor in the perception of human-likeness.

Li (2006) *did* find a significant negative correlation between backchannel quantity and conversation participant's self-reported enjoyment of an interaction in a cross-cultural dialogue between Anglo-Canadians and Mainland Chinese speaking in English. This study differs from Study 1 and Poppe et. al. (2011) in that the study was not asking a

third party to rate their perception of a quality of the listener in the conversation; however, the significant result of this study compared to the non-significant results of Study 1 and Poppe et. al. raises the question of if quantity of backchannels might have an effect during intercultural communication, but not have an effect during intracultural communication.

Differences in the research goals and designs of Study 1 and Poppe et. al. (2011) could be responsible for the differences in data of each study. Overall, the goal and experimental design of Poppe et. al. differed greatly from the goal and design of Study 1. Poppe et. al's research aimed to gather information to aid in the development of a human-like artificial listener. The goal of Study 1 was to learn if manipulating backchannel quantity has an effect on the perception of the listener, specifically in regard to listening competence. Besides the differences in research questions or goals, the two studies had very different experimental designs and stimuli. While in both of my experimental studies, I exclusively looked at verbal backchannels, Poppe et. al. were interested in both verbal as well as non-verbal backchannels. Thus, Poppe et. al. presented participants with visual stimuli, a side by side view of an actual speaker and an animated version of an artificial listener. Participants of this study simply had different information provided to them than participants of my experimental studies who only had audio stimuli.

The fact that Poppe et. al.'s (2011) rate of 6 to 12 backchannels was based on both verbal and non-verbal backchannels could be an explanation as to why the data from Study 1 were non-significant If counting both verbal and non-verbal backchannels, a

larger number (like for example, nine) might be perceived more positively, but if only considering verbal backchannels, that larger number might not transfer over.

Another possible explanation for why the results of Study 1 were non-significant could be that the quality of human-like and the quality of good listening (listener competence) are not comparable. For this study, I assumed that high listening competence is a salient feature of what makes someone (or some*thing*, in the case of an artificial agent) human-like; however, this was motivated only by intuition. Each characteristic might have a different ideal backchannel quantity range.

The goal for Study 2 was to learn if manipulating backchannel placement impacts the perception of listening competence of the person uttering backchannels. It was hypothesized that manipulating the location of backchannels would have an effect on the perception of listening competence. It was additionally hypothesized that the clause-break condition would have the highest rating for listening competence (Dittman & Llewelyn, 1967; Ward & Tsukahara, 2000; Cathcart et. al., 2003; Kjellmer, 2009), followed by the noun-phrase condition, with finally the word-medial condition having the lowest rating for listening competence (Kjellmer, 2009).

No results from Study 2 were significant. Poppe et. al. (2011) manipulated what they referred to as *timing* of backchannels in their study, which for all intents and purposes is the same as backchannel placement. Poppe et. al. *did* find a manipulation of backchannel timing to have a significant effect on the perception of the human-likeness of the listener. The independent variables used in their study were much more general

than the ones that I used in Study 2. Their variables were *original* timing (leaving the timings of backchannels that a real listener had used during the conversation intact) and *random* timing (simply inserting backchannels randomly into the conversation). The more general variables used in Poppe et. al. could account for the differences in the results from each study. Additionally, as described earlier when discussing Study 1, human-likeness and listening competence could be two characteristics that cannot be compared, thus differences in data between them would be expected.

Although the results from Study 2 were not significant, they were in the hypothesized direction. Given the amount of research that has shown that in American English, backchannels occur most frequently at clause breaks (Dittman & Llewelyn, 1967; Ward & Tsukahara, 2000; Kjellmer, 2009), it makes sense that the listener perceived as most competent would be the one who produced backchannels at clause breaks.

3.3.2 Limitations

The Perceptibility of Backchannels. While backchannels are an important part of everyday discourse, they might not be a particularly salient feature of conversation in the minds of speakers. Campbell-Kibler (2005) explains that "sociolinguistic variation is structured in such a way that certain variables and variants are more perceptible to speakers and listeners than others, leading to different patterns of use and different likelihoods of taking on social meaning" (p. 8). In her work, she uses the example of how speakers and listeners are more conscious of variances in *-ing* ending differences than

they are of differences between different pronunciations of the /t/ phoneme. There is a possibility that like the /t/ phoneme, speakers and listeners might not be particularly perceptive of differences in quantity and placement of backchannel, leading to the inconclusive results of both studies. The choice of independent variables might have also played a role. There is a possibility that the independent variables *I* chose to use as in this research were too similar for differences to be noticeable by participants.

Choice of Independent Variables. For Study 1, the three independent variables were 3 backchannels, 9 backchannels, and 15 backchannels. There is a possibility that the differences in these values were simply too small for there to return significant results. Future research using my methodology could use independent variables that were more different from each other: for example, a condition with *no* backchannels, and a condition with 30 or more backchannels.

For Study 2, it is again possible that the independent variables chosen were not different enough from each other to return significant results. Given the fact that the backchannel insertion points were chosen based off data from English (Kjellmer, 2009), it is possible that all of the insertion points were familiar to participants, leading to little difference in the results. Given the fact that differences in backchanneling location exist between different languages, such as between Arabic and English (Ward, et. al. 2007), one possible method of approaching choosing backchannel insertion points in the future could be by referencing insertion points used in other languages. For example, having one condition with backchannels inserted at the most common location in English (clause

breaks), another condition with backchannels inserted according to the most common location in Arabic, etc. Since the results for Study 2 were inconclusive but were in the hypothesized direction, it would be beneficial for future research to explore what the results would be if the variables chosen for backchannel location were more dramatically different.

Operationalization of *Competent Listening.* The dependent variables used in both studies (*understanding*, *(pays) attention,* and *friendliness)* were how I chose to operationalize *competent listening* based primarily off Bodie et. al. (2012) for the purposes of these experiments. Bodie et. al.'s (2012) original framework of listening outlines five salient features of good listening. The two that were not included in the dependent variables for these studies were *responsiveness* and *conversational flow* (i.e. how well the listener helped to maintain conversational flow in their interaction). Expanding the number of dependent variables by including these two other features from Bodie et. al. (2012) could impact future results due to information about more properties of good listening would be collected.

Another option for future operationalization of listening competence is to operationalize it more in line with past sociolinguistic research on the perception of the speaker. Feldstein, Dohm, & Crown (2001) combined the adjective ratings of *intelligent*, *confident*, *competent*, and *ambitious* to create the dependent variable score of *perceived competence* of a speaker in an experiment manipulating the gender and speech rate of a speaker. Earlier research specifically concerned with the perception of the listener used

very similar qualities to create measures of competence. Brown et. al. (1973) combined rated results on the adjectives *intelligent, confident, ambitious, active,* and *good-looking*. Street et. al. (1983) used *intelligent, confident, ambitious, sincere, good-looking,* and *effective communicator*.

The current study did not operationalize *competent listening* in a way similar to how the competence of a speaker had been operationalized in past research. Future research that investigates the relationship between backchanneling and listening competence could approach operationalizing *competent listening* using adjectives such as *intelligent, confident,* and *competent,* despite these adjectives not having a specific link to listening, as the adjectives that were chosen from Bodie et. al. (2012) do.

3.4 Summary of Findings

The results from the two experimental studies were all inconclusive. I did not find that the manipulation of backchannel quantity or placement has an effect on the perception of the listener. Nevertheless, these results provide important empirical data and raise several theoretical and methodological questions that I discuss in the next chapter.

Chapter 4: Methodological and Theoretical Contributions

This chapter discusses the methodological and theoretical contributions of this thesis. To summarize, the findings of this thesis are the following:

1. My research offers a new experimental methodology for testing the relationship between backchannels and the perception of the listener.
2. Theoretically, this research contributes to the area of perception studies conducted in sociolinguistics. This is an area of study that traditionally studies the perception of the speaker in an interaction. This thesis expands on this domain by researching the perception of the *listener*.
3. This thesis contributes experimental data about what the ideal number and placement of backchannels are for a *competent listener*.
4. Finally, this thesis raises theoretical questions about how the concept of a *competent listener* should be operationalized.

4.1 Methodological Contributions

This thesis offers a new experimental methodology used for testing the relationship between backchannels and the perception of the listener in a conversation. This is an important contribution given the fact that the relationship between backchannels and listening is one that lacks experimental research describing it. The need for more quantitative research methods in the field of linguistics has been brought to focus by Fedorenko & Gibson (2010) and Gibson & Fedorenko (2013). My novel

methodology falls in line with their call for more quantitative research methods in linguistics.

Fedorenko & Gibson (2010) and Gibson & Fedorenko (2013) highlight the lack of quantitative research methods in the subfields of syntax and semantics; however, I believe that this methodological weakness exists throughout other fields of linguistics as well. The lack of experimental research on the relationship between backchanneling and listening can be attributed to the weak methodological standards brought up by Fedorenko & Gibson and Gibson & Fedorenko. The research by Fries (1952) and Yngve (1970) resembles the methods critiqued by Gibson & Fedorenko (2010). Their methods consist of one researcher (the author) evaluating and making judgments off a small number of examples. Although this kind of analysis can suffice for initial explorations of a linguistic phenomenon, it does not allow for scientific hypothesis testing, nor does it supply quantitative data.

Additionally, there has been a collection of research done on backchanneling, such as Stubbe (1998), Ward et. al. (2007), and Poppe et. al. (2011), that while experimentally sound, were developed based on descriptions from those such as Fries and Yngve, which had *not* been explored experimentally. Thus, by not having information collected experimentally on the original relationship between backchannels, subsequent research is built on a "shaky foundation," at least from the experimental perspective.

What is the benefit of conducting experimental research on a subject, such as the relationship between backchanneling and listening, for which qualitative research already

exists? Although qualitative research is very useful within the field of linguistics, researchers coming from more quantitatively rigorous fields might ignore these descriptions given their lack of scientific methodologies. It is important that research in linguistics rises to meet the high standards of fields that place strong importance on experimentally derived data (Gibson and Fedorenko, 2013). Not only to improve the credibility of linguistic research to other disciplines but also to provide other fields of research with quantitative data that might be of use to them. In an increasingly data-fueled world, experimentally collected findings about backchannels and the perception of the listener could positively impact areas that might need data on this topic for their work, such as for developing teaching materials for L2 learners of English on backchanneling or optimizing AI and virtual agent's abilities to communicate with people.

My research provides a framework to experimentally examine the claim made by Fries (1952) that backchannels are used "to signal that [the listener] is listening attentively to the speaker" (p. 50), by investigating the relationship between backchanneling and the perceived listening competence of the listener. These experimental methods can be used by future researchers to further explore the connection between backchanneling and listening.

4.2 Theoretical Contributions

Traditionally, research investigating the perception of a conversation participant has focused on how the *speaker* is perceived (Labov, 1966; Labov, 1972; Kramer, 1977;

Feldstein, Dohm, & Crown, 2001; Campbell-Kibler, 2005). My research has expanded on perception research from sociolinguistics by investigating the perception of the *listener*.

The next theoretical contribution that has come out of this research is experimental data on the ideal quantity and placement of backchannels associated with a *competent listener*. Since the results of both experimental studies were inconclusive, this research was not able to pin down exactly what the ideal quantity and placement of backchannels for a competent listener are. Nevertheless, the data is still useful to researchers interested in this topic and provides a foundation for future research aiming to get statistically significant data regarding these questions.

While there has been literature that explores how manipulating the quantity (Li et. al., 2010; Poppe et. al., 2011) and placement (Poppe et. al., 2011) backchannels impacts the perception of other interactional variables (such as overall enjoyment of the conversation or *human-likeness* of the listener), there is a lack of literature that explores the specific theoretical question of how a manipulation of backchannel quantity and placement impacts the perception of the listener. My thesis addresses this question experimentally. The inconclusive results of study 1 suggest that the ideal quantities of backchanneling for another social feature (how human-like a listener is) are possibly not applicable to what the ideal quantity of backchannels are for a person to be deemed a competent listener. Although the results of Study 2 were inconclusive, the fact that the results were in the hypothesized direction suggests that clause breaks have the potential to be the ideal location of backchannels for a person to be deemed a competent listener.

This aligns with previous literature that shows clause breaks to be the most common insertion point for backchannels in American English (Dittman & Llewelyn, 1967; Ward & Tsukahara, 2000; Kjellmer, 2009).

Finally, this thesis has raised questions about how the concept of a *competent listener* (referred to in some literature as a *good listener*) should be operationalized. While there had been previous literature that operationalized the concept of a *competent speaker* (Brown et. al., 1973; Street et. al., 1983; Feldstein et. al., 2001), there had not been any research done on operationalizing the concept of a *competent listener.* Given the importance listeners play in didactic conversations literature on perception of the listener would benefit from having an operationalization of the concept of *competent listening.* Although this thesis might not have perfected the operationalization of the concept (given the non-significant results of the experimental studies), this research did begin to lay the ground for operationalizing the concept of a good listener and most importantly, this research has identified the need to operationalize the concept of a *competent listener*.

What is the use of researching the perception of the listener at all? Additionally, what is the use of operationalizing, and through that, having a better understanding of the concept of a *competent listener*? Conversations are a collaborative effort between people in which the roles of speaker and listener are passed off between participants. If there is only a solid understanding of the concept of a *competent speaker* and how speakers are perceived, we only have information about half of the story (or in this case, half of the conversation!). By collecting more data about the perception of the speaker and having a

better understanding of what it means to be a *competent listener*, we gain a better understanding of conversations and communication overall.

Chapter 5: Conclusion

5.1 Summary

The present thesis investigated the relationship between backchanneling and the perception of the listener through two experimental studies. In these studies, the quantity and placement of backchannels were manipulated. The level of how much of a *competent listener* participants of the studies perceived the backchanneler to be was measured.

It was initially hypothesized that manipulating both the quantity and placement of backchannels would have an effect on the perception of the backchanneler's listening competence; however, there were no significant differences in the results of both studies. Despite the inconclusive results of the experimental studies, this research makes both methodological and theoretical contributions.

While the relationship between backchannels and listening has been described qualitatively in theoretical research and been explored in fields like discourse and conversational analysis, there has been a lack of experimental research about this relationship. Theoretically, this thesis has three main contributions. Firstly, while perception studies in sociolinguistics have traditionally focused on the perception of the speaker, this research expands on this tradition of research by investigating perceptions of the listener. Secondly, this research contributes data and analysis about what the ideal number and placement of backchannels are for a *competent listener*. Finally, this thesis brings into discussion the question of how a *competent listener* should be operationalized.

This research operationalized a *competent listener* using three features. A *competent listener* is someone who is *understanding* the speaker, *paying attention* to the speaker, and is *friendly* towards the speaker (Bodie et. al., 2012). Given the inconclusive results of the experimental studies, this operationalization of a *competent listener* might not be the most effective, highlighting the need for future research that focuses on how this concept should be operationalized.

5.2 Directions for Future Research

In order to gather more data regarding the ideal quantity and placement of backchannels, future research should conduct studies that recreate the methodology provided in this thesis using different independent variables. These independent variables would be chosen to be more dramatically different from one another. For example, a recreation of Study 1 would use a wider range of different quantities while a recreation of Study 2 could use backchannel insertion points from different languages as experimental conditions. This would be done to increase the likelihood that there be a significant difference between the results of the conditions.

Another venue for future research is to experimentally explore what it means to be a *competent listener*. Research exploring this subject could choose to follow the methodologies of research that examine the perception of competent speakers (Brown et. al., 1973; Street et. al., 1983; Feldstein et. al., 2001). Alternatively, future research could develop its own methodology for investigating this question.

References

Allen, R., & Brown, K. (1976). *Developing communication competence in children.* Skokie, Ill.: National Textbook

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*(6), 941–952.

Bennett M, & J. Jarvis. (1991). The Communicative Function of Minimal Responses in Everyday Conversation. *Journal of Social Psychology*, *131*(4), 519–523.

Beňuš, Š., Gravano, A., & Hirschberg, J. (2007). Prosody of backchannels in American English. In *Proceedings of 16th International Congress of Phonetic Sciences* (pp. 1065-1068).

Bodie, G. D., St. Cyr, K., Pence, M., Rold, M., & Honeycutt, J. (2012). Listening Competence in Initial Interactions I: Distinguishing Between What Listening Is and What Listeners Do. *International Journal of Listening*, *26*(1), 1–28.

Brown, B. L., Strong, W. J., & Rencher, A. C. (1972). Manipulation of vocal qualities by speech synthesis: A new way to study person perception. In *Proceedings of the Annual Convention of the American Psychological Association*. American Psychological Association.

Cathcart, N., Carletta, J., & Klein, E. (2003). A shallow model of backchannel continuers in spoken dialogue. In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics - EACL '03* (Vol. 1, p. 51).

Campbell-Kibler, K. (2006). *Listener perceptions of sociolinguistic variables: The case of (ING).* (Unpublished doctoral dissertation). Stanford University. Stanford, CA.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Dittman, A. T., & Llewellyn, L. G. (1967). The Phonemic Clause as a Unit of Speech Decoding. *Journal of Personality and Social Psychology*, *6*(3), 341–349.

Deng, X. (2013). Listener response. In J. V. Sigurd D hondt, Jan-Ola Östman (Ed.), *Handbook of Pragmatics* (pp. 1–25). John Benjamins Publishing.

Fedorenko, E., & Gibson, E. (2010). Weak quantitative standards in linguistics research. *Trends in Cognitive Science*, *14*(6), 233–234.

Feldstein, S., Dohm, F. A., & Crown, C. L. (2001). Gender and speech rate in the perception of competence and social attractiveness. *Journal of Social Psychology*, *141*(6), 785–806.

Fries, C. (1952). *The structure of English: An introduction to the construction of English sentences*. New York: Harcourt, Brace.

Gibson, E., & Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes*, *28*(1–2), 88–124.

Hymes, D. (1972). On communicative competence. *Sociolinguistics: Selected Readings*, 269–293.

Jackendoff, R. (1999). Possible Stages in the Evolution of Language Capacity. *Trends in Cognitive Science*, *3*(7), 272–279.

Jurafsky, D., Shriberg, E., Fox, B., & Curl, T.S. (1998). Lexical, Prosodic, And Syntactic Cues For Dialog Acts. In *Proceedings of ACL/COLING-98 Workshop on Discourse Relations and Discourse Markers,* 114–120.

Kramer, C. (1977). Perceptions of Female and Male Speech. *Language and Speech, 20*(2), 151-161.

Kjellmer, G. (2009). Where do we backchannel?: On the use of mm, mhm, uh huh and such like. *International Journal of Corpus Linguistics*, *14*(1), 81–112.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology, 4*(3), 343-346.

Kraut, R. E., Lewis, S. H., & Swezey, L. W. (1982). Listener responsiveness and the coordination of conversation. *Journal of Personality and Social Psychology*.

Labov, W. (1966). *The social stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular* (University of Pennsylvania Press conduct and communication series ; 3). Philadelphia: University of Pennsylvania Press.

Li, H. Z. (2006). Backchannel Responses as Misleading Feedback in Intercultural Discourse. *Journal of Intercultural Communication Research*.

Li, H. Z., Cui MEd, Y., Wang Ph, Z. D., Prufer Leske, I., & Aguilera, L. (2010). *Backchannel Responses and Enjoyment of the Conversation: The More Does Not Necessarily Mean the Better. International Journal of Psychological Studies* (Vol. 2).

Poppe, R., Truong, K. P., & Heylen, D. (2011). Backchannels: Quantity, Type and Timing Matters. In K. R. T. Hannes Högni Vilhjálmsson, Stefan Kopp, Stacy Marsella (Ed.), *Intelligent Virtual Agents* (pp. 228–239). Reykjavik, Iceland, : Springer.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh'and other things that come between sentences. *Analyzing discourse: Text and talk*, *71*, 93. Schiffrin, D. (1987). *Discourse Markers*. Cambridge University Press.

Street Jr, R. L., Brady, R. M., & Putman, W. B. (1983). The influence of speech rate stereotypes and rate similarity or listeners' evaluations of speakers. *Journal of Language and Social Psychology*, *2*(1), 37-56.

Stubbe, M. (1998). Are you listening? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics, 29*(3), 257-289.

Tolins, J., & Fox Tree, J. E. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics*, *70*, 152–164.

Ward, N. G., Escalante, R., Al Bayyari, Y., & Solorio, T. (2007). Learning to show you're listening. *Computer Assisted Language Learning*, *20*(4), 385-407.

Ward, N., & Tsukahara, W. (2000). *Prosodic features which cue back-channel responses in English and Japanese. Journal of Pragmatics* (Vol. 32).

Yngve, V. H. (1970). On Getting A Word In Edgewise. *Papers from the Sixth Regional Meeting [of the] Chicago Linguistic Society*, 567–578.

Appendix A: Transcripts of Recordings, Study 1

3-Backchannel Condition

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

Narrator: Ugh, well, first I woke up and tried to take a shower, but the hot water was out so I had to take a cold shower and it just made me so cranky. But like, whatever, it happens all the time so I'm kind of used to it but it's not any less any less uh any less annoying. But why does it only ever happen when it's really cold out, you know?

Listener: Uh-huh

Narrator: Like why does the hot water never go out when it's like, 90 degrees out? Okay so then, the bus to campus was so crowded. I just, ugh, wish there were more buses running during rush-hour! It makes *no* sense why there aren't. Anyways, the bus is *jammed* and I barely squeezed my arm through to hold on to a pole so I didn't like, fall over. And, I was next to this guy who had just god-awful B.O.

Listener: Uh-huh

Narrator: It was so unpleasant. Okay okay, and *then*, when I got to campus I was all stressed out, so I decided to treat myself to a coffee, to try and like, cheer myself up or something and when I finally got my coffee I spilled it all over myself! Ugh! So I had to go to the bathroom to get myself cleaned up and then that made me late for class. So Yeah, I'm not in a great mood today.

9-Backchannel Condition

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

Narrator: Ugh, well, first I woke up and tried to take a shower, but the hot water was out so I had to take a cold shower

Listener: Uh-huh

Narrator: and it just made me so cranky. But like, whatever, it happens all the time **so** I'm kind of used to it

Listener: Uh-huh

Narrator: but it's not any less any less uh any less annoying. But why does it only ever happen when it's really cold out, you know?

Listener: Uh-huh

Narrator: Like why does the hot water never go out when it's like, 90 degrees out? Okay so then, the bus to campus was so crowded.

Listener: Uh-huh

Narrator: I just, ugh, wish there were more busses running during rush hour! It makes *no* sense why there aren't. Anyways, the bus is *jammed* and

Listener: Uh-huh

Narrator: I barely squeezed my arm through to hold on to a pole so I didn't like, fall over. And, I was next to this guy

Listener: Uh-huh

Narrator: who had just god-awful B.O. It was so unpleasant. Okay okay, and *then*, when I got to campus I was all stressed out so I decided to treat myself to a coffee,

Listener: Uh-huh

Narrator: to try and like, cheer myself up or something and when I finally got my coffee I spilled it all over myself! Ugh!

Listener: Uh-huh

Narrator: So I had to go to the bathroom to get myself cleaned up and then that made me late for class. So Yeah, I'm not in a great mood today.

Listener: Uh-huh

15-Backchannel Condition

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

Narrator: Ugh, well, first I woke up and tried to take a shower, but the hot water was out

Listener: Uh-huh

Narrator so I had to take a cold shower and it just made me so cranky.

Listener: Uh-huh

Narrator: But like, whatever, it happens all the time so I'm kind of used to it

Listener: Uh-huh

Narrator: but it's not any less uh any less annoying. But why does it only ever happen when it's really cold out,

Listener: Uh-huh

Narrator: you know? Like why does the hot water never go out when it's like, 90 degrees out?

Listener: Uh-huh

Narrator: Okay so then, the bus to campus was so crowded. I just, ugh, wish there were more buses running during rush-hour!

Listener: Uh-huh

Narrator: It makes *no* sense why there aren't. Anyways, the bus is *jammed*

Listener: Uh-huh

Narrator: and I barely squeezed my arm through to hold onto a pole

Listener: Uh-huh

Narrator: so I didn't like, fall over. And, I was next to this guy

Listener: Uh-huh

Narrator: who had just god-awful B.O. It was so unpleasant.

Listener: Uh-huh

Narrator: Okay okay, and *then*, when I got to campus I was all stressed out **so** I decided to treat myself to a coffee,

Listener: Uh-huh

Narrator: to try and like, cheer myself up or something and

Listener: Uh-huh

Narrator: when I finally got my coffee I spilled it all over myself! Ugh!

Listener: Uh-huh

Narrator: So I had to go to the bathroom to get myself cleaned up

Listener: Uh-huh

Narrator: and then that made me late for class. So I'm not in a great mood today

Listener: Uh-huh

Appendix B: Transcripts of Recordings, Study 2

Clause-Break Condition

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

Narrator: Ugh, well, first I woke up and tried to take a shower, but the hot water was out so I had to take a cold shower

Listener: Uh-huh

Narrator: and it just made me so cranky. But like, whatever, it happens all the time so I'm kind of used to it

Listener: Uh-huh

Narrator: but it's not any less any less uh any less annoying. But why does it only ever happen when it's really cold out, you know?

Listener: Uh-huh

Narrator: Like why does the hot water never go out when it's like, 90 degrees out? Okay so then, the bus to campus was so crowded.

Listener: Uh-huh

Narrator: I just, ugh, wish there were more buses running during rush-hour! It makes *no* sense why there aren't. Anyways, the bus is *jammed* and

Listener: Uh-huh

Narrator: I barely squeezed my arm through to hold onto a pole so I didn't like, fall over. And, I was next to this guy

Listener: Uh-huh

Narrator: who had just god-awful B.O. It was so unpleasant. Okay okay, and *then*, when I got to campus I was all stressed out so I decided to treat myself to a coffee,

Listener: Uh-huh

Narrator: to try and like, cheer myself up or something and when I finally got my coffee I spilled it all over myself! Ugh!

Listener: Uh-huh

Narrator: So I had to go to the bathroom to get myself cleaned up and then that made me late for class. So Yeah, I'm not in a great mood today.

Listener: Uh-huh

Noun Phrase Condition

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

Narrator: Ugh, well, first I woke up and tried to take a

Listener: Uh-huh

Narrator: shower, but the hot water was out so I had to take a cold

Listener: Uh-huh

Narrator: shower and it just made me so cranky. But like, whatever, it happens all the time so I'm kind of used to it but it's not any less uh any less annoying. But why does it only ever happen when it's really cold out, you know? Like why does the hot

Listener: Uh-huh

Narrator: water never go out when it's like, 90 degrees out? Okay so then, the bus to campus was so crowded. I just, ugh, wish there were more

Listener: Uh-huh

Narrator: buses running during rush-hour! It makes *no* sense why there aren't. Anyways, the bus is *jammed* and I barely squeezed me

Listener: Uh-huh

Narrator: arm through to hold onto a pole so I didn't like, fall over. And, I was next to this

Listener: Uh-huh

Narrator: guy who had just god awful B.O. It was so unpleasant. Okay okay, and *then*, when I got to campus I was all stressed out so I decided to treat myself to a coffee, to try and like, cheer myself up or something and when I finally got my

Listener: Uh-huh

Narrator: coffee I spilled it all over myself! Ugh! So I had to go to the bathroom to get myself cleaned up and then that made me late for class. So yeah, I'm not in a great

Listener: Uh-huh

Narrator: mood today

<u>Word-Medial Condition</u>

Backchannel locations are marked with a bolded arrow ↑

Listener: Hey Lisa!

Narrator: Oh, hey Jen! How are you?

Listener: I'm okay, you?

Narrator: Ugh, I've been better, I've just had the *worst* morning.

Listener: Yeah? What happened?

> Ugh, **we↑ll**, first I woke up and tried to take a shower, but the hot water was out so I had to take a cold **sho↑wer** and it just made me so cranky. But like, whatever, it happens all the time so I'm kind of used to it but it's not any **le↑ss** any less uh any less annoying. But why does it only ever happen when it's really

cold out, you know? Like why does the hot water **ne↑ver** go out when it's like, 90 degrees out? Okay so then, the bus to campus was so crowded. I just, ugh, wish there were more buses **run↑ning** during rush-hour! It makes *no* sense why there aren't. Anyways, the bus is *jammed* and I barely squeezed my arm through to hold onto a **po↑le** so I didn't like, fall over. And, I was next to this guy who had just god awful B.O. It was so unpleasant. Okay okay, and *then*, when I got to cam↑pus I was all stressed out so I decided to treat myself to a coffee, to try and like, cheer myself up or **some↑thing** and when I finally got my coffee I spilled it all over myself! Ugh! So I had to go to the **bath↑room** to get myself cleaned up and then that made me late for class. So Yeah, I'm not in a great mood **to↑day**

Appendix C: Survey Materials for Studies 1 & 2

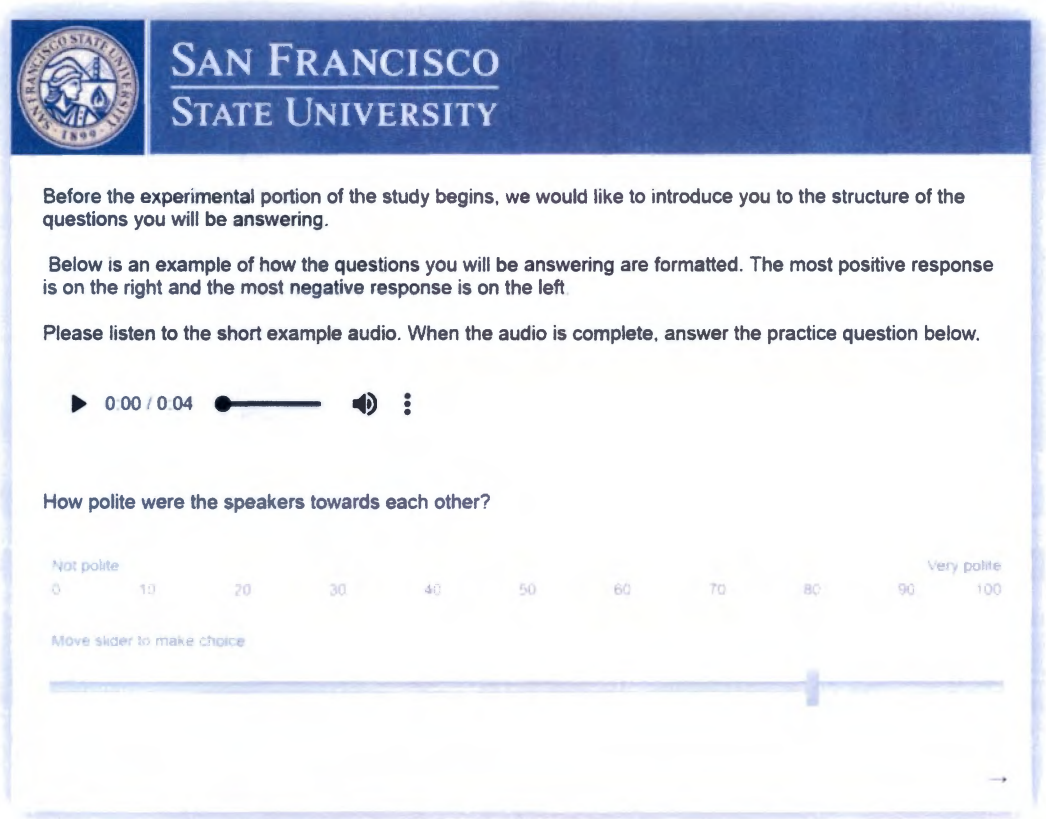

Figure C.1: Training Page

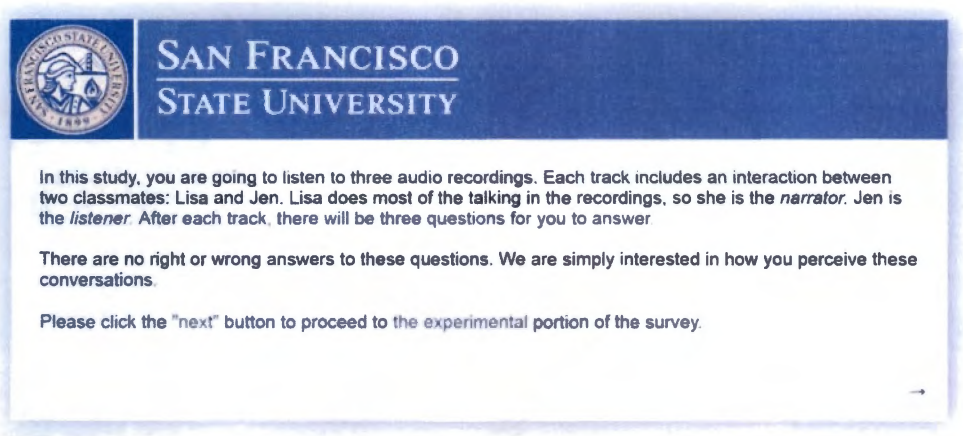

Figure C.2: Introduction Page

Figure C.3: Survey Page

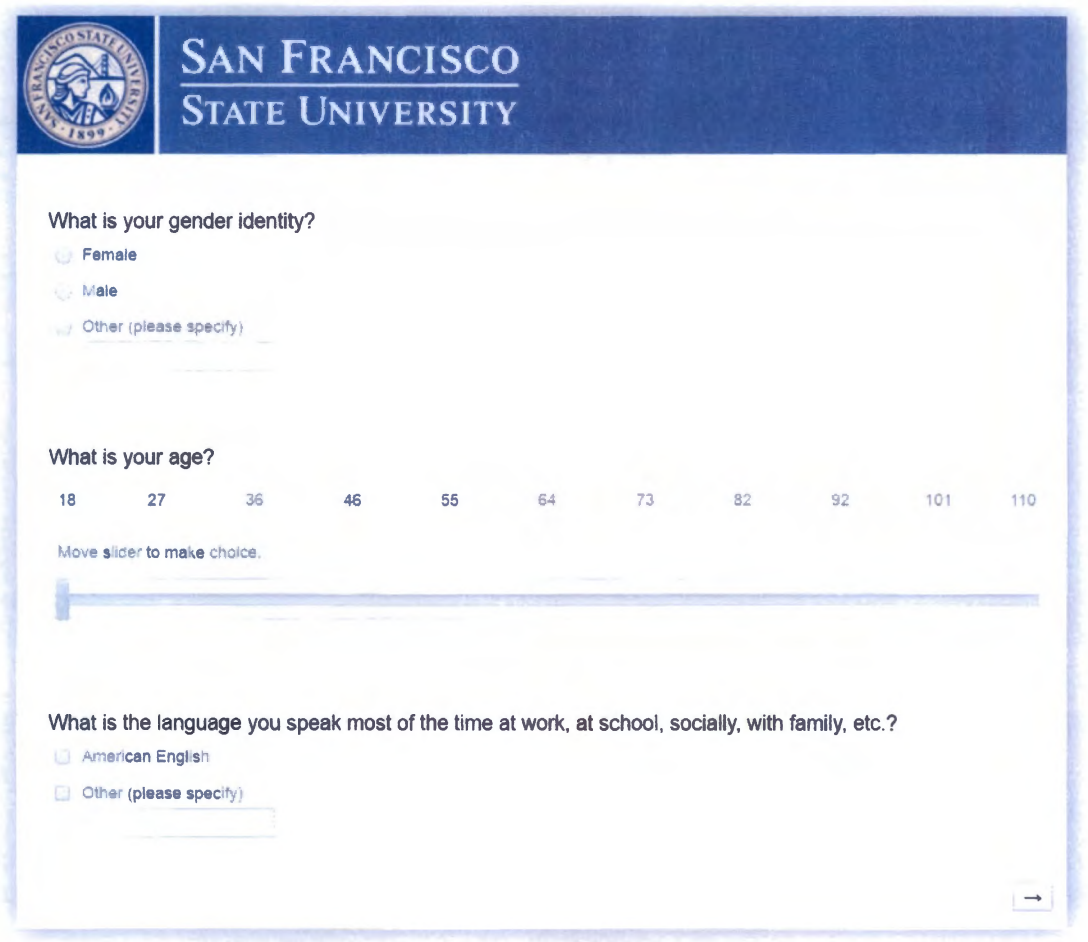

Figure C.4: Demographics Page