2012-12-01

The Effect of Pause Duration on Intelligibility of Non-Native Spontaneous Oral Discourse

Ryan Frederick Lege
Brigham Young University - Provo

The Effect of Pause Duration on Intelligibility

of Non-Native Spontaneous Oral Discourse

Ryan F. Lege

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Mark W. Tanner, Chair
Dirk A. Elzinga
Wendy B. Smemoe

Department of Linguistics and English Language

Brigham Young University

November 2012

ABSTRACT

The Effect of Pause Duration on Intelligibility
of Non-Native Spontaneous Oral Discourse

Ryan Lege
Department of Linguistics, BYU
Master of Arts in TESOL

Pausing is a natural part of human speech. Pausing is used to segment speech, negotiate meaning, and allow for breathing. In oral speech, pausing, along with other suprasegmental features, plays a critical role in creating meaning as comprehensible speech is seen as a goal for language learners around the world. In order to be comprehensible, language learners need to learn to pause correctly in their speaking. Though this notion is widely accepted by applied linguists and many language teachers, the effect of pausing on intelligibility of spontaneous oral discourse has not been established by empirical data. This study isolates pause duration in spontaneous oral discourse in order to establish its connection to the intelligibility of non-native speech. In this study, North American undergraduate students' reactions to non-native pause duration in spontaneous oral discourse were examined. The task involved measuring the NESs' processing, comprehension, and evaluation of three different versions of an international teaching assistant's presentation: One with unmodified pause duration, one with pause duration shortened by 50%, and a third passage with pause duration lengthened by 50%. Results showed a positive correlation between pause duration and number of listeners able to identify main ideas. Finally, listener reaction was measurably more positive to the unmodified passage than to the passages with lengthened or shortened pauses.

ACKNOWLEDGEMENTS

Table of Contents

List of Tables

Chapter 1

Introduction

Pauses play a critical role in the communication of oral discourse (see Chambers, 1997, p. 538). Pauses allow the speaker time to take a breath. They direct the listener's attention to important words or phrases in the discourse. They also function similarly to punctuation in written language to group words into constituents and meaningful units/phrases (Griffiths, 1991; Reid, 2003). In the pronunciation pedagogy literature, these meaningful pieces or units are termed "thought groups" (Gilbert, 1993).   Segmenting speech into appropriate thought groups is particularly important for second language (L2) learners. In order to be understood, L2 learners must segment their speech by using pauses for maximum intelligibility (Chen, 1996). Intelligibility is defined by Derwing and Munro (2005) as "the extent to which the speaker's intended utterance is actually understood by a listener" (p. 385). Intelligibility refers to a listener's actual comprehension of a stream of speech.

Because of the importance of pausing, many researchers advocate its inclusion in pedagogical practice (see Gilbert 1983). Chen sums up pausing's importance and its connection with pedagogy in this manner: "Students need to know where to pause in sentences when saying or reading a passage aloud. If pauses are not put in the right places, the listener will have difficulty processing the meaning of the sentences, even if the individual words are pronounced well" (1996, p. 12). Therefore, pausing plays a central role in the negotiation of meaning from discourse.

In the time period from the mid-1980s to the present, L2 pronunciation instruction has seen a marked shift in focus toward an emphasis on prosodic features (stress, pitch, linking) of which pausing is a key component (see Levis & Grant, 2003, p. 13). Levis (2005) indicates that

this trend in language teaching for more than two decades has been to emphasize suprasegmentals over segmentals has been due in large part as a means to promote intelligibility (p. 369). Empirical research over the past 15 years has begun to confirm that suprasegmental instruction does indeed promote greater degrees of intelligibility on the part of native speakers listening to non-native speech (Derwing, Munro, & Wiebe, 1997; Derwing & Rossiter, 2003). In 2005, Derwing and Munro, two researchers who have conducted numerous studies dealing with accent, pronunciation instruction, and intelligibility, have called for more empirical research in an effort to better understand how suprasegmentals influence learners' speech and the comprehension of that speech. They cite Hahn's 2004 research investigating primary stress as precisely the type of work needed to confirm or refute approaches advocated by pronunciation specialists which emphasize suprasegmental instruction. Hahn's study investigated primary stress as a particular feature influencing intelligibility. Empirical data gathered in her study showed that correctly implemented primary stress led to higher intelligibility ratings and more favorable evaluative ratings. The study described in this paper employs a similar research design to Hahn's study, with the focus on another prosodic feature, that of pause duration in spontaneous speech and its effect on intelligibility.

Chapter 2

Review of Literature

The Nature of Pausing

Pausing has been a subject of interest in various contexts for a number of years. Scientists and researchers in many fields such as psychology, pedagogy, and linguistics, have examined the nature and definition of pauses in speech. Duez, a linguist, supplied the definition that a "pause is defined as an interval of the oscillographic trace, where the amplitude is indistinguishable from that of the background noise" (1982, p. 13). By this definition, a pause is evident at any period during which sound wave oscillations are not present. Analysis of any segment of speech will clearly show the presence of pauses. By this definition, however, each time amplitude cannot be separated from background noise, a pause would be identified. A study investigating pausing would also need to include in its definition of a pause, the length of the gap or duration that constitutes a pause. Otherwise, short gaps between lexical units would therefore qualify as pauses, making analysis difficult to carry out. A threshold needs to be defined in order to differentiate between small absences of sound and pauses that occur in the speech stream. In Griffith's 1991 research of speech rates in language classrooms, an absence of sound for 100 milliseconds was considered as a pause. A gap of 100 milliseconds is a very sensitive threshold. This pause length includes even the most minuet space between individual lexical units. Additionally, in Griffith's study, pauses were identified by ear perception, then analyzed, indicating a use of subjectivity which does not lend itself to establishing a standard. However, Riaszantseva's (1999) study of second language proficiency and pausing also included measurements of pauses beginning with a duration of 100 milliseconds (ms) and extending to a length of three seconds. Riaszantseva references Goldman-Eisler (1961) as her rational for a

100ms cut-off point, adding that 100 ms is sufficiently low for capturing all possible pauses that may influence language proficiency ratings (p. 39). Other researchers, however, have adopted considerably higher cut-off points. Oliveira (2002), for example, adopted a minimum cut-off point of 250 ms, stating that "since 71.5% of all pauses occur in the duration interval between 250 ms and one second." He utilized a "standard" set by the various works of Goldman-Eisler in the 1960s (2002, p. 539). Interestingly enough, both of the values, 100 ms and 250 ms, proposed by Oliveira (2002) and Riazantseva (1999) were based on the same works of Goldman-Eisler. More recently, Derwing, Munro, Thompson, and Rossiter (2009) have argued that a cut of point of 400 milliseconds has been established as "common practice", using their prior studies and review of literature as a rationale (p. 542). Derwing and Munro's 2004 study of second language fluency adopted a 400 millisecond cut off point to measure unfilled pauses following the value set by Riggenbach's 1991 study (2004, p. 665). Riggenbach (1991) explored the role of hesitation phenomena in determining the overall fluency of English language learners. Based on her review of literature, Riggenbach utilized a hesitation classification hierarchy to rate hesitations found in her data. Citing the works of Deese (1980) and Fillmore (1979), Riggenbach notes that pauses below 400 ms "are considered 'micropauses' in the literature on fluency (p. 426). A micropause would be considered a normal part of fluent speech. However, Riggenbach also classified hesitations as "a silence of .3 to .4 seconds" (p.426), and included these hesitations in her analysis. Though researchers have used different values for determining what constitutes a pause, from .25 seconds to .4 seconds, the literature does not indicate that researchers have come to any consensus as to what is considered the minimum and maximum length of a pause.

Pausing is often viewed as a factor of disfluent speech (see Rossiter, 2009, p. 398); however, pausing is not an uncommon or wholly negative feature of fluent language. Pauses are utilized as space for breathing and thinking when participating in any form of oral discourse (see Griffiths, 1991). Pauses also play other, normal important roles in human interaction; Chambers (1997) states, "In verbal encounters common to all language users, mainly conversations and discussions, pauses and hesitations are normal features of interaction" (p. 538). Pausing is not only a normal feature of speech interaction, but according to Richards (1983), "in natural speech, between 30% and 50% of speaking time may consist of pauses and hesitations" (p. 225). He then goes on to say that pausing is indicative of "selection and planning processes speakers make use of" (p. 225). Pausing is part of the natural process of negotiating and creating meaning in natural fluent spontaneous speech. Riazantseva (2001) succinctly summarizes pausing's role in speech production in this manner:

> Results of L1 studies on pausing have suggested that the apparent intermittent quality and smoothness of speech are not present in what the speaker actually says. Instead, periods of silence constitute an indispensable part of the speech production process, the stage of speaking activity at which planning of the speech occurs. (p. 497)

Pausing is natural for all fluent speakers of all languages, and is not limited to learners of a language. Lennon (1990) adds that "indeed, pause is itself a necessary ingredient in fluent speech" (p. 408). Additionally, pausing plays an integral role in outlining the semantic and pragmatic meaning of an utterance. Chafe (1985) beautifully characterizes the relationship of pausing and speech in this manner: "[Speaking] is a creative act, relating two media, thought and language, which are not isomorphic but require adjustments and readjustments to each other" (p. 78). The nature of speaking, which requires these types of reformulations and readjustments,

indicates that pausing is a natural part of the negotiation of thought and meaning in speaking. Additionally, pauses intrinsically are manifest in regular, logical locations. Chambers (1997) further explicates where pauses occur by saying, "Natural pauses, allowing breathing space, usually occur at some clause junctures or after groups of words for a semantic unit" (p. 539). The logical location of pauses in the speech stream suggests that pausing is in fact a systemized process of segmenting speech.

Being a natural part of spoken discourse, pauses are present in all language from that of the politician to that of the pre-literate. Pausing is a natural phenomenon that serves the important task of separating speech into thought groups. Gilbert (1993) highlights the importance of pauses in natural speech by equating their importance with other suprasegmental features in this manner, "pauses are nearly as important as the correct stress pattern of a word or correct emphasis on sentence focus" (p. 41). By creating thought groups, pauses in effect draw attention to important information in a sentence which is emphasized by other suprasegmental features. Prator (1985) states that "Within thought groups, words and syllables are not pronounced as separate units; they flow along smoothly, without jerkiness, and one seems to blend into the next. A person who did not know any English would find it hard to tell where one word ended and another began" (p. 34). English's system of reductions, linking, and varied syllabic weight is difficult to decode and comprehend, but coupled with appropriate pausing it is easier for listeners to digest. In addition, Andrade (2005) further comments on the importance and role of thought groups by addressing them from a learner perspective. Andrade conducted a study on Japanese university students to ascertain the difficultly of learning different suprasegmental features for this particular group of learners. In his rational for including thought groups as a major focus of his study, he comments that "one of the major perception problems

for learners is the task of segmenting the stream of speech into these meaningful subunits, a process known as segmentation and thought grouping" (2005, p. 7). Andrade and other researchers who have studied thought groups indicate that pausing plays a crucial role in creating meaning from discourse.

Besides describing the nature of pauses and their positions, other research has sought to clarify what constitutes natural pausing. In particular, Ruder and Jensen (1972) investigated the nature of pause duration in different sentence structures. They used five recorded sentences within which the words *lost* and *contact* were manipulated to vary in complexity by sentence. Twelve raters mechanically adjusted the pauses between the two tokens, *lost* and *contact*, to locate the minimum fluent pause duration and the optimal pause duration for the particular sentence. Ruder and Jensen found that hesitation pause duration "did differ significantly as a function of syntactic structure, in that hesitation pauses perceived within a subordinate clause boundary were significantly longer in duration than were those occurring in other types of syntactic boundaries" (1972, p. 49). However, the degree of variability between raters was considerable, meriting further investigation into what constitutes optimal pause duration. Nevertheless, the results of this study indicate that an increase in linguistic complexity merits longer pauses, giving the speaker time to process and produce the structure. Duez (1982) confirms that the use of pausing varies greatly depending on the type and complexity of discourse. More complex types of discourse, requiring a higher register and addressing abstract topics, will necessitate more pauses and pauses of greater duration. Oliveira's 2002 research has also confirmed this finding. Oliveira looked at the role of pausing within an oral narrative framework. The narrative framework separated the speaking process into six different sections, beginning from an abstract (introduction) and ending with a coda (conclusion). He found that

depending on the function and complexity of the particular section of the narrative, the pause to speech ratio varied (p. 540). Oliveira concludes that pause to speech ratio "varies as a function of task complexity; people tend to pause more often and to remain in silence for much longer when performing more complex task[s]" (p. 541). Indeed speakers make regular use of pausing to segment information into digestible chunks of language.

History of Pausing Research

Linguistics, psychologists, and teachers all recognize the importance of pausing in language used in social contexts. Since the 1950s, researchers have published various descriptive studies examining the nature of pauses and their role in speech. In the 1960s, linguists and other scientists began to define and classify pauses into various types (see Goldman-Eisler, 1961). A great example of this is Barik's 1968 classification of pauses into two types: juncture and hesitation pauses. He asserted that juncture pauses divide the speech stream into digestible segments, while hesitation phenomena occur when the speaker is unsure of the content (p. 156). These terms remain widely used today when referring to and classifying different types of pauses.

In the language teaching and learning field in the 1980s, researchers began to investigate the role of pauses in the language acquisition process. A central component of this research was the influence of pauses on the concept of fluency. During the 1970s, innovative language teaching and learning methods not only directed attention to the learner, but further established the notion of language as a vehicle for communication. This communicative focus meant that greater importance was attributed to the notion of fluency. With a focus on fluency, research on pausing as a factor impacting fluency expanded (see Macias, 2006; Raupach, 1987; Riggenbach, 1991; Rossiter, 2009; Simensen, 2010).

Fluency and pausing. Language teachers and learners have long struggled to accurately define the term fluency and its role and place in the language learning process. A typical definition of fluency is given by the *Cambridge Encyclopedia of Language*, "smooth, rapid, effortless use of language" (Crystal, 1987, p. 421). English language teachers often use similar components to define fluency: speed and a lack of frequent pauses. Therefore, in conventional definitions, pausing has a prominent position as a detractor from overall fluency.

Riggenbach (1991), a researcher who examined the nature of pausing, suggests that there are two types of pauses (or hesitations), silent gap and non-lexical fillers (p. 438). Riggenbach conducted a study using these two types of pauses within discourse to determine which type had the greatest effect on overall fluency. Based on her data, she came to the conclusion that "the frequency of unfilled pauses is a strong indicator of non-fluency" (p. 441). Non-lexical fillers are, according to Riggenbach, more unnatural and indicative of a lack of competency in the target language. Riggenbach also took her analysis even further by examining the impact of pauses on comprehensibility. Riggenbach found that clusters of more than 3 hesitations have a negative effect on comprehensibility (see Chambers, 1997, p. 540). The nature of, and environment in which the pauses occur, greatly impacts the comprehensibly of oral discourse. Building on the work of Riggenbach, Macias (2006) specifically investigated the connection of silent pauses to fluency. She comments that "silent pauses are defined as any interruption in the flow of speech which is manifested in silent form, independently of the motive for the pause or the point at which it is produced" (p. 28). Macias' study used simultaneous interpretation as a tool for examining the role of pauses in determining fluency. Macias chose to examine pauses of length greater than or equal to two seconds, considering this duration to be "unnatural" interruptions of the speech stream (pauses of duration .25 to 2 seconds were considered natural,

Pauses were isolated based on listener perception. He validated perception of the pauses through use of additional raters. The results of this study indicate that silent pauses are indeed related to mean ratings of fluency.

Other research has confirmed the role of pauses as determinants of overall fluency, even in languages other than English. Raupach (1987) researched the concept of overall fluency by examining the language of exchange students before and after their experience in France. The students were rated on their overall fluency by native speakers of French. Raupach analyzed the features of the student's language that led to their being labeled on a continuum from disfluent to fluent. Raupach then analyzed the linguistic features present in the language of both the fluent and disfluent students to ascertain which features contributed to overall fluency. Raupach came to this conclusion at the end of his research, "fluency cannot be reduced to speed of delivery, and articulation rate has l*ess impact on our perceptions of fluency than the length, the nature, and the location of pauses in an utterance*" [italics added] (p. 150).

Raupach's observation not only suggests that pausing is the most important feature of fluency, but he also illuminates some key elements of pausing to consider: length, nature, and location. Correct implementation of these three characteristics leads to fluency in the target language. Additionally Chambers (1997) states that, "Becoming fluent…is not about speaking faster (articulation rate), but about pausing less often and pausing at the appropriate junctures in an utterance" (p. 540). Pausing plays a critical role in fluently using a language. Indeed, Simensen (2010) researched characteristics of fluent speech and found that "characterizations such as 'unhesitating flow', 'spontaneity' and 'few noticeably long pauses' are used as positive features of language at upper proficiency" (p. 8). High-proficiency English learners utilize pauses of the appropriate length to make their language more comprehensible. Without the

proper amount and length of pausing, learner-language is perceived as unnecessarily rapid. Anderson-Hsieh and Koehler (1988) reported that one of the most common complaints about non-native speech is that "it is too fast to understand" (p. 561). Rate of speech may not be the greatest contributing factor to this perception; incorrect pausing may be the source of this perception. Towell's 1987 and 1977 studies also validate the conclusion that pausing is in fact the most important element of fluency. Towell tested students before and after studying abroad in another country. Towell compared the students' speech samples to determine what factors were contributing to an increase in fluency. He concludes that "the main factor contributing to an improvement in fluency was identified as an increase in the mean length of runs (i.e. the speech unit between two unfilled pauses)" (Towell, 1996, p. 84). It is clear through the research that pausing is in fact the most critical factor in determining fluency.

Rossiter (2009) conducted a study providing empirical data that confirms the connection between pausing and fluency. Raters in Rossiter's study included a heterogeneous group of ESL learners (n =24) from a variety of language backgrounds, the "speakers" and "listeners," a group which included non-native English speakers (NNSs) with extensive ESL teaching experience (and a master's degree or more), native English speaking (NS) university students, and advanced NNS. Speakers performed a picture-narration task, to which the "listeners" made fluency judgments. Listeners recorded their impressions about the speakers and assigned ratings of fluency to the speakers. Speakers were rated twice based on samples pre and post 10 week ESL instruction. Rossiter calculated the correlation of pausing to the overall fluency rating given to the audio passages and found all values within the range of -0.625 to -0.728 statistically significant (p. 404). As the frequency of pauses increased, perception of fluency decreased. She further found that pausing accounted for "almost half of these [negative temporal] impressions"

(p. 407). All raters, both expert and non-expert, similarly found pausing to be an important negative factor, detracting from overall fluency.

While pausing has been demonstrated to have an influence on fluency, pausing itself can be examined through many different methods. The effect of pausing on communication can be evaluated by investigating pause length, the nature of pauses, the location of pauses, pause frequency, and pause duration, the latter of which is the focus of this study. Although pause duration has been addressed as an independent variable, its effect on fluency has not been firmly established. Kormos and Denes (2004) found issue with the treatment of pause duration in the literature, claiming the following:

> Most of these studies, however, suffer from several methodological shortcomings partly because they use very few subjects and fail to employ reliable methods of analysing the duration of pauses (e.g. Ejzenberg, 2000; Freed, 1995; 2000; Lennon, 1990; Riggenbach, 1991). In many of them, no statistical analyses and computer technology for identifying pauses reliably were used. (p. 147)

Though prior studies have established a tenable connection between pause duration and fluency, Kormos and Denes' (2004) research employs objective measures of pause identification to add additional validity to their study. Kormos and Denes explored the role of pause duration in predicting "native and non-native speaking teachers' perception of fluency" (p. 147). Their research involved skilled teachers assigning fluency ratings to speech samples containing pauses of differing lengths. Kormos and Denes then explored the reasons why some samples received higher fluency ratings than others. Their results confirm findings reported in earlier studies, such as Riggenbach (1991). Kormos and Denes report that "advanced students spoke faster, silent pauses took up a smaller proportion of their speaking time, produced longer stretches of

discourse between pauses, used shorter pauses and uttered more stressed words within a minute than low-intermediate students" (p. 157). Interestingly enough, their discussion emphasized the importance of increasing the frequency of stressed words per minute to promote fluency. This conclusion reinforces the importance of thought groups, which all contain one prominent syllable and are bounded by pauses.

Many studies have focused on the role of pausing and hesitation phenomena in defining and describing fluency. Fluency is obviously an important way to examine the ability of a language learner. Interestingly enough, the term fluency is often used synonymously with proficiency in a language. Indeed, when the term fluency is applied to a language learner's interlanguage, it suggests a high degree of communicative competence (Chambers, 1997, p. 536). Fluency is connected to how well a speaker is understood, as well as how interlocutors perceive the speaker (see Rossiter, 2009, p. 396).

Intelligibility

Due to the connection between fluency and understanding a speaker, researchers have attempted to establish an association between the suprasegmental features that influence perception of fluency with the concepts of intelligibility and comprehensibility. Though intelligibility and comprehensibility are often used interchangeably in the literature with various nuances in meaning, this study will use the term intelligibility as defined by Derwing and Munro (2005) to be "the extent to which the speaker's intended utterance is actually understood by a listener" (p. 385). By defining intelligibility as what is actually understood by the speaker, Derwing and Munro suggest that intelligibility is synonymous with the degree to which a learner *comprehends* or understands the speech stream. Comprehensibility, also in accordance with a definition offered by Derwing and Munro (2005), is used to refer to listener's perception of the

intelligibility of the speaker. In other words, comprehensibility measures a listener's perception of how easy it is to understand a speaker. This concept has been typically measured though Likert scale questions.

Derwing, Munro, and Thomson (2004) investigated the connection between intelligibility, pausing, and fluency. Their study investigated the nature of fluency and its connection to various prosodic elements. Their study also supports the establishment of a connection between fluency and comprehension. One of their research questions is of particular interest to this discussion: "What are the relations among fluency and both comprehensibility and accentedness?" (2004, p. 661). In order to investigate this question, Mandarin speakers completed a speaking task in response to a visual stimulus. Next, they answered a question about the happiest moment in their lives. Finally, the Mandarin speakers were then asked to question the researchers about the researcher's happiest moment. After completing these speaking tasks, the audio recordings of the Mandarin speakers were randomly presented to undergraduate raters. These raters listened to randomized speech samples and rated the speakers on fluency, accentedness and comprehensibility. In their analysis, the researchers also analyzed the pauses, speech rate, and MLR (or Mean Length of Run) between pauses. They also found pausing to have a significant correlation with fluency ($r = .731$, p <.02) (p. 667). Next, they examined comprehensibility. They found that comprehensibility was significantly correlated with fluency judgments for all tasks (p. 669). They conclude that "our data provided an indication that listeners' perceptions of comprehensibility have a clearer tie to fluency judgments than do their assessments of accentedness" (p. 673). This suggests that speech rate and pausing are more important than the ability of a speaker to produce each individual phoneme of the language. The concept of fluency also relies heavily on the degree to which the person who makes judgments on fluency

understands the speaker's utterances. As a crucial factor in determining overall fluency, logically, pausing must hold importance in the intelligibility of the speech.

If asked about their language learning goals, many English language learners would likely respond that native-like pronunciation is one of these. However, language teachers would likely assert that intelligibility is a far more realistic goal. In choosing intelligibility as a language learning goal, the pronunciation teacher is inclined to ask which language features have the most influence on intelligibility. Intelligibility is the target that pronunciation instruction traditionally sets for itself, but English language teachers know little about how best to help learners achieve it.

A number of revealing studies have attempted to trace correlations between various features of nonnative speech and native speakers' intelligibility judgments. Anderson-Hsieh and Koehler (1988) reported that "prosodic deviance may affect comprehension more adversely than does segmental deviance" (p. 562). In a related study, Derwing, Munro, and Wiebe (1998) studied the effects of both segmental and suprasegmental instruction on learners' comprehensibility ratings and concluded that the latter had a greater effect on performance in communicative contexts. These two studies in particular illustrate the importance of focusing on suprasegmentals such as pausing. Results such as these helped to reconceptualize the idea of pronunciation teaching to include features such as pausing.

Attention to pausing as an important element within the suprasegmental canon has been furthered by research by Fayer and Krasinski (1995). Fayer and Krasinski are some of the only researchers to confront pausing and intelligibility as manifested in spontaneous speech. Their study was conducted as a follow up to research completed in 1987, which indicated hesitations were the most detrimental factor towards comprehension of the overall message (1995, p. 116).

The 1995 study improved upon the design of the prior study by allowing for more empirical measurement. The results revealed a relatively high correlation between intelligibility rankings and total pause time and length of the longest pause (p. 119). Fayer and Krasinski make an interesting conclusion. They state, "It appears that as long as pauses are deemed appropriate in length, even though they might be frequent, they are not distracting to the listener" (p. 119). In their words, even if pauses are frequent, pause duration, if maintained at appropriate levels allows the listener to decode the speech stream adequately. Additionally, this conclusion suggests that there must be an acceptable range of pause duration, or optimal pause duration, which increases the overall comprehensibility of speakers.

As a result of studies such as these and the anecdotal proliferation of suprasegmental instruction, Levis (2005) indicates that the trend of pronunciation teaching "[f]or the past 25 years,...ha[s] emphasized suprasegmentals [features which include intonation, stress, rhythm, and pausing] rather than segmentals in promoting intelligibility" (p. 369). The reasoning for such a trend is a belief that suprasegmental instruction is an accelerated path to intelligibility. Field (2005) comments, "The notion of intelligibility has become central to the teaching of pronunciation" (p. 369). While suprasegmental instruction has been lauded and implemented as a path to success, it has been pointed out that there is a lack of empirical evidence to support these claims. In 1992, Anderson-Hsieh, Johnson, and Koehler argued that prosody "provides the framework for utterances and directs the listener's attention to information", but that these conclusions "are based on informal observations rather than on empirical evidence from research studies" (p. 531). Since 1992, many studies have focused on prosodic elements. However, many of these studies have focused on other prosodic elements in languages other than English, descriptive linguistics, or psychology (see De Pijper, 1994; Fujisaki, 2004; White, 1997).

In addition, a number of studies have focused on prosody in general, rather than focusing on a single suprasegmental feature and its effect on intelligibility. Field's 2005 study of the effect of lexical stress on intelligibility and Hahn's 2004 article (based on her 1999 dissertation work) focused on primary stress are examples of studies that have attempted to "isolate particular suprasegmental features for analysis" (Hahn, 2004, p. 201). Researchers have supported a connection between pausing and comprehensibility, but have provided little empirical data to support their claims aside from general data about suprasegmentals and comprehensibility. Still, studies emphasize the importance of pausing as a critical feature for generating meaning through spoken language. De Pijper's 1994 study is an example of a study that explores the connection of pausing and comprehensibility. De Pijper focused his study on the strength of boundaries as perceived by listeners of Dutch. His rationale for conducting the study, which compares the features of spoken and written language, is enlightening in the manner that it highlights the importance of pausing in communicating a speaker's message. De Pijper (1994) comments that, "Only some of the most important boundaries between groups of words are shown by punctuation marks such as commas and periods. In speech, a speaker uses suprasegmental means to indicate the degree to which he or she feels words are grouped together, or, to take a different perspective, separated" (p. 2037). This statement is particularly enlightening if we view it from the interlocutor's perspective. Without the means of suprasegmental cues, the listener will not be able to adequately segment the speech and extract meaning. In other words, by this trail of logic, the speech will not be comprehensible. Additionally, De Pijper goes on to extol the virtues of prosodically proper speech in this manner: "If synthetic speech is provided with appropriate prosody, it becomes more natural, more pleasant to listen to and linguistically more transparent" (p. 2037). According to De Pijper, even our overall perception of speech is altered by

suprasegmental implementation. Language is easier to decode when supplemented with prosodic cues. However, prosody is a decidedly multi-faceted feature of speech. The following statement by De Pijper (1994) offers insight into the most important prosodic features: "Of all possible suprasegmental cues a speaker may use to convey prosodic structuring of utterances, it is generally maintained that position and length of pauses, pitch and pre-final lengthening are the most important ones" (p. 2038). Pauses, when placed correctly, and held for the appropriate time period, segment and organize speech for uptake by the listener. The results of De Pijper's study also indicated that longer pauses over 300ms are associated with higher perceptual boundary strength (PBS) values (p. 2043). This finding supports the idea that 300 ms is a natural cut-off point for pauses that are used to segment speech into understandable chunks. De Pijper's study provides a look into what makes speech comprehensible. However, there is an important caveat. His study was conducted with Dutch as the primary language. Though pausing is a language universal, length and placement may vary language by language, meaning that results could be different if focusing on the English language.

Other researchers have confirmed that pausing has an impact on the comprehensibility of spoken language. Blau (1990) compared the comprehension of a passage played at a normal rate, a slowed rate, and a lengthened rate in which the passage had long three second pauses inserted at sentence, phrase and clause boundaries. She found that "pauses at constituent boundaries, on the other hand, do enhance the comprehensibility of aural input significantly" (p. 752). Due to the length of the pauses, the overall rate of delivery was reduced significantly from approximately 200 words per minute to 150 words per minute. Overall rate of speech is a possible threat to the validity of the study, but Blau admirably included an additional moderating variable of speech rate to investigate its impact on comprehensibility. She included two

additional variations of her passage, one mechanically slowed and the other sped up. Her results showed that "the version with pauses did in fact yield significantly higher comprehension than both other versions" (p. 749). This study makes a valuable contribution to our knowledge of pausing; nevertheless there are some limitations that call for further research. First, Blau's passage was an oral reading, not spontaneous speech. Duez (1982) indicates that these are two completely different types of discourse that contain different pausing patterns. Second, three second pauses at phrase, clause, and sentence boundaries are unnatural in normal human speech. Few studies attempt to test pausing's effect on comprehensibility in spontaneous speech through empirical means. This is an area which calls for further research and study.

Unlike Blau, who sought to find the effect of lengthening pauses in speech, Reid (2003) digitally manipulated speech to see if pauses could be shortened without negatively impacting comprehension. Reid used a portion of the Brown, Carlsen, Carstens (BCC) Test of Listening Comprehension as the instrument. The test includes five sections that test immediate recall, following directions, lecture comprehension, recognizing transitions, and recognizing word meanings. His analysis of the audio passages used on the test indicated that the normal pause length between clauses was between 300ms and about 500ms on average. Using digital means, he prepared four versions of this text: one with the original pausing intact, one in which all pauses were shortened by 300ms, one in which all pauses were shortened by 150ms, and one in which all the pauses were lengthened by 300ms. Reid justifies these particular lengths by referring to recommendations from past studies (p. 32). To investigate the impact of pauses, Reid played the different versions of the passage for four research groups (n =144, comprised of intact classes). Students were tested on their listening comprehension by completing the test tasks. Reid found that the passage with the longest pauses, lengthened by 300ms, had a

significant positive effect on the raters' ability to recognize transitions, but he found that the control group performed significantly better than the other groups at recognizing word meanings. Overall, Reid was unable to demonstrate a significant correlation between pauses and comprehensibility. He cites some concerns with the face validity of the lecture comprehension portion of the test (p. 63), which may not have been truly measuring comprehension of the lecture, in turn making it difficult to establish a viable connection between pausing and comprehensibility. Reid's study highlights some of the difficulties in establishing this connection, but does not discount a correlation. Other studies though, such as the 2004 study by Derwing et al have established a tenable connection by demonstrating a correlation between comprehension and fluency judgments. This study indicated that in various speaking tasks, pruned syllables and pauses were significantly correlated with fluency ratings, and in turn with intelligibility ratings.

Chen (1996) also investigated the relationship of intelligibility and fluency in his study of the factors that contribute to Chinese learners' unique challenges with learning English. Chen specifically focused on the elements influencing Chinese learners of English. One of the key variables in his study was pausing. He comments on the relation of pauses, intelligibility, and instruction in this manner: "Students need to know where to pause in sentences when saying or reading a passage aloud. If pauses are not put in the right places, the listener will have difficultly processing the meaning of the sentences, even if the individual words are pronounced well" (p. 12). Chen's analysis singles out pausing as a critical feature of the English language, and hints at its importance though there is a lack of empirical data to support the claim. Likewise, Chambers (1997) describes the effect of pauses placed in incorrect positions, illuminating pausing's importance but failing to provide more than expert opinions to support the claim. Chambers

concludes that "pauses appearing at places other than these are judged as hesitations, revealing either lexical or morphological uncertainty" (p.539). This finding provides insight into the nature of pausing; these hesitations are crucial to our understanding of how pausing impacts our understanding of speech. Nevertheless, Chambers research provides an expert opinion supporting the notion of pausing, but doesn't add critical empirical evidence. The difficult nature of pausing, a multifaceted prosodic feature consisting of placement and duration, creates difficulties producing empirical data.

An additional difficulty arises when considering the nature of pauses. Goldman Eisler, in a 1968 publication, noted that pause duration varies with individual speakers (p. 11). This knowledge complicates looking at pauses and attributing numerical values to natural pause duration. When studying the effect of pausing on the intelligibility of speech, we need to look at acceptable ranges of optimal pause duration, rather than create arbitrary standards leading to simple binary judgments on the pertinence of pause durations. Krivokapic (2004) performed an interesting study on pause duration which accounted for this problematic individual variation. Though multiple speakers often concur about the positions of pauses at sentence junctures, variations in pause placement occur more frequently in sentence-internal positions. With this in mind, Krivokapic (2004) attempted to normalize the pausing in her subjects' speech to appropriately analyze these pauses (p. 164). To create more standardized discourse for analysis, Krivokapic used a synchronous reading paradigm to offset pause variations in different idiolects. Subjects were native speakers who read aloud sentences together with a partner, reducing the variability of pause duration in the speech produced. Thus, together the speakers pause for more standard durations and in logical positions, drawing on their shared knowledge of linguistic timing (p. 166). Krivokapic, analyzing the pauses of the synchronous readings, found that longer

phrases had significantly longer (preceding and following) pauses (p. 167). This research provides some insight into correlation between pauses and intelligibility. More complex, longer segments of speech require longer pauses to enable uptake. It should be noted that this study involved pausing as manifested in oral readings, not spontaneous speech. Pausing in spontaneous speech is more complex and difficult to isolate and subjectively measure in a research study.

Andrade (2005) performed a study examining different suprasegmental features impact on non-native speakers' comprehension of oral speech. A portion of Andrade's study on Japanese college students focused on the importance of properly understanding thought groups in oral discourse. Raters in his study were presented with a test paper consisting of pairs of phonemically identical sentences, which varied only due to the existence of "absence of a pause, link, juncture, or other feature that changed the meaning" (2005, p. 11). Raters listened to a recording and answered a question, indicating their understanding of the proper thought group. Andrade reports that this section of the test was among the most difficult sections, but found it difficult to interpret the results of the section. He calls for more testing of a similar type which includes a greater quantity and variety of test items. His research suggests that thought groups are among the most difficult suprasegmental features for learners to identify and understand. Future work needs to be done to confirm Andrade's results on a larger and more generalizable scale.

Shortcomings notwithstanding, researchers have widely accepted that there must be connections between pausing and intelligibility, but in order to justify these claims in the scientific world, there needs to be well designed empirical studies that yield good data. This study attempts to rectify this issue by addressing pausing's effect on intelligibility. Pausing is

fairly complex as a prosodic feature, which includes both the placement and duration of each pause. Fayer and Krasinski's 1995 research study indicated that pause duration had a greater effect on intelligibility than pause frequency (p. 118). To test the question of pause duration, an empirical study would need to investigate how pause duration and not pause placement or frequency influences intelligibility. To test the influence of pause duration, this feature would have to be isolated within the discourse for analysis. In order to properly isolate pausing for analysis, this study attempts to replicate a study aimed at isolating the feature of primary stress.

Hahn's 1999 dissertation, the results of which have been published succinctly in 2004, is exemplary due to its attempt to isolate a particular suprasegmental feature. Hahn's 2004 research focused on "the impact that NNSs' misplaced or missing primary stress could have on native-speaking listeners" (p. 205). She also wished to determine whether speakers with correct primary stress placement were evaluated more favorably by native speakers (p. 206). Hahn recorded a non-native English speaker performing an oral reading of a short academic passage. From this reading, she then proceeded to create three distinct versions of the passage: The first having correctly placed primary stress, the second having misplaced primary stress, and the third lacking placement of primary stress. Hahn found that subjects who listened to the sample with correctly placed primary stress better recalled details and information from the lecture. When primary stress was placed correctly, Hahn also found that "subjects generally responded more positively to NNS discourse" (p. 213). Hahn's study confirmed through empirical data that the suprasegmental feature of primary stress does in fact have a significant impact on listener comprehension. Hahn's study concluded with a call that "similar studies of other linguistic variables...would help us to more fully understand intelligibility and NS reactions to nonnative pronunciation" (p. 218). Derwing and Munro (2005) praised Hahn's work as "precisely the type

of work needed to confirm or refute approaches emphasizing suprasegmentals," and added that "teachers' goal should be to analyze students' speech and help them select areas for practice based on empirical findings" (p. 387).

In accordance with this call to action, the study reported in this paper examines the effect of pause duration on the intelligibility of non-native spontaneous speech through empirical data collected by employing methodology similar to that used in Hahn's study. The research questions investigated in this study are as follows:

- What effect does the duration of pausing at sentence and clausal boundaries within spontaneous oral discourse have on the intelligibility of non-native speech by native English listeners?

- Is there an optimal pause duration for maximum intelligibility in non-native speech?

- What effect does pause duration have on evaluative reactions of the non-native speaker?

Chapter 3

Research Design

The purpose of this study was to measure the effect of pause duration on native English speakers' comprehension of non-native English spontaneous oral discourse. To fulfill this purpose, this study employed research methods similar to those used by Hahn in her 1999 dissertation work. In her study, she isolated the feature of primary stress to study its effect on the comprehension of oral discourse. The current study, however, focused on researching a different prosodic feature, that of pause duration. The goal of employing similar methodology as Hahn was to gather empirical data about pause duration and its role in communication. This chapter will provide detailed information about the research design used, including the participants (raters), the instrument used, and the procedure that was followed.

Participants

Listeners. A total of 95 native-English speaking undergraduate students from Brigham Young University's College of Humanities participated as raters for the study. Of the 95 total raters, 32 were randomly assigned to listen to a passage containing unmodified pauses, 32 to a passage with lengthened pauses, and 31 to a passage with shortened pauses. Raters were selected from the College of Humanities so as to reduce their level of familiarity with the topic of the passage used in the listening task.

Speaker. The speaker used for the audio recording was an advanced non-native English speaker whose L1 is Japanese. At the time of recording, he was in his second year of a Masters in Food Science at Brigham Young University. The speaker had been studying at a university level in the United States for a period of about 6 years. He was selected from a pool of three highly proficient graduate students (two Japanese, one Korean) within the college of Life

Sciences at Brigham Young University. All three were recorded speaking about their prospective majors and lecturing on specific topics within their fields. For this study, the speaker was particularly important, because he would need to meet three basic criteria. First, in order to test native speaker's comprehension and reactions to the speech of a non-native speaker, the speech sample needed to have certain accented features that would identify the speaker as a non-native English speaker. Second, the speaker's language would need to be free from any glaring segmental or suprasegmental errors that could possibly act as confounding variables in the study. Third, the preliminary speech samples needed to exhibit natural pausing patterns and durations. A Japanese speaker who was completing his final year of a degree in Food Science while acting as a teaching assistant in university-level classes was chosen as the final candidate for creation of the passage used in the study. Once a variety of samples were recorded, the researcher began analysis of the speaker's language. After a careful analysis of the various samples, the researcher confirmed that the speaker's language met all three criteria.

Instrument

Audio passage. The features of the audio passage were vitally important to the validity of the study, so time and attention were given to producing the best possible passage. Creating the passage involved cautious topic selection so as to limit raters' potential familiarity with the theme. Varonis and Gass' research on familiarity with topic (1985) indicated that "familiarity with a particular task decreases but does not totally compensate for the need to negotiate meaning" (p. 160). Their research, involving the use of information gap activities, indicates that if interlocutors possess a degree of background knowledge, their comprehension will be enhanced in a predictable and measurable manner. The researcher and speaker worked together to choose a topic that would be a basic introduction to some aspect of Food Science. The

speaker was first recorded explaining a variety of topics. He was allowed to create a simple outline with key words to scaffold his presentation, but prohibited from writing complete sentences, which he would be tempted to read as he spoke. Avoiding oral reading ensured that the audio text coincided with the goal of the study: To assess the intelligibility of a non-native speaker's spoken language, not intelligibility of an oral reading. Before additional work was carried out, a panel of four TESOL and Linguistics professionals met together and listened to samples of the speaker's speech. Following this, members of this panel discussed segmental and suprasegmental features of the speaker's language and reviewed the goals of the study. The panel reached a consensus that the speaker's language would in fact be appropriate for the study in that there was no obvious deviance in segmental usage that would detract listeners from focusing on the message of a produced passage.

Following the initial analysis, the researcher and speaker endeavored to create a passage to use for the study. Together with the speaker, the researcher chose a basic topic related to Food Science. A topic was chosen that would be unfamiliar to the uninitiated in Food Science, but still simple and relevant enough to produce an engaging passage. The topic chosen focused on the purpose of food science and the basic nutrients that comprise foods. To accomplish the goal of measuring the effect of pause duration on intelligibility of speech, a target passage length of two to three minutes was chosen by the researcher as an appropriate length for the audio sample. Many researchers have chosen different sample lengths to study comprehension of a listening passage. This study employed a passage of approximately three minutes in duration based on commentary and results of previous research. The literature indicates that 20 seconds is a sufficient minimum value for making judgments on intelligibility. Derwing, Munroe, Thomson and Rossiter (2009) offer this insight as to their rationale for choosing 20 second samples by

stating, "There is no consensus in the field concerning the appropriate length of speech samples for fluency evaluation by judges. In earlier studies, we determined that 20 seconds was sufficient for raters to make a reliable assessment" (p. 541). On the other hand, Fayer and Krasinski (1995) chose two minutes as an appropriate length for judging intelligibility. They state that "approximately two minutes of speech per speaker were sufficient for listeners to be able to ascertain judgments" (Fayer & Krasinki, 1995, p. 116). However, their decision to employ two minute samples seems based on their own reasoning and rationale. Derwing (2009) argues that "Using substantially longer speech samples not only lengthens the duration of the experiment, which in turn increases the demands on listeners, but also has the potential to make the rating of individual samples more difficult, because the raters must hold in memory a longer narrative." (p. 541). Though they reason that 20 second samples are sufficient, they also concede that trade-offs were made in this regard, though they neglect to mention specific concerns (p. 541). Laura Hahn's 2004 study of the effect of primary stress on comprehension used a 269 word sample which, read at a normal rate, greatly exceeds 20 seconds, and instead falls around the two minute mark. Laura Hahn's study was effective at determining the effect of stress on intelligibility, so we can assume this method was successful. In order to provide enough data for analysis and abide by the principles of listening task creation, the researcher chose to use a three minute sample of speech. Three minutes provided sufficient time for the speaker to share basic elements about the topic of food science and provided enough text for raters to adequately indicate their comprehension of the passage. This length also allowed for the isolation of particular pauses by segmenting the passage into chunks of language within which the effect of individual pauses could be analyzed.

With a passage of this target length in mind, the speaker was recorded in a recording studio on Shure microphones in 16-bit AIFF audio using a PowerMac computer. Once recorded, the passage was checked to make sure it conformed to principles of listening comprehension task creation. Buck (2001) states, "Texts with slower speech rates tend to be easier than texts with faster speech rates" (p. 149). Lass and Prater (1973) investigated listener preferences with regards to speech rate. In their literature review, they note that previous studies have shown values of 163 to 175 wpm to be preferred by listeners (p. 96). Using various speech rates applied without distortion of the pitch or quality of the audio signal to two different speech types, oral reading and impromptu speech, they found that a rate of 175 wpm was most preferred in both oral reading and impromptu speech (p. 100). The speaker of the passage for this study on pausing and intelligibility spoke at an appropriate rate of 153 wpm, which Lass identified among the top three preferred speech rates for spontaneous speech (p. 100).

Summarizing Anderson and Lynch (1988), Buck (2001) comments that "Texts with longer pauses between idea units tend to be easier than texts with shorter pauses between idea units, or no pauses at all" (p. 149). Buck (2001) also reports that Freedle and Kostin (1996, 1999) examined 337 TOEFL Multiple choice items to determine what made items difficult. They created an inventory of characteristics that made items more difficult. Among these items, the most important contributing factor was the location of the necessary information. "When the necessary information came near the beginning of the text, or when it was repeated, the item tended to be easier" (Buck, 2001, p. 153). In accordance with this finding, the speaker added a brief introduction and example that would not be used to test intelligibility of the passage.

Finally, when it was determined that the passage possessed characteristics necessary to complete the study (see Appendix A for a transcript of the passage), an in depth analysis of

pausing in the passage was conducted using Praat, an open-source language analysis tool (Boersma & Weenink, 2012), to identify the pauses within the passage, which were then each manually verified by comparing Praat's identification to the audio waveform. As stated in the review of literature, pauses above 250 ms were identified as unfilled pauses. Therefore, pauses were identified at or above a 250 ms, 300 ms, and 400 ms cut-off point (see Derwing, et al., 2009; Oliveira, 2002; as well as Chapter One of this document to see a discussion of cut-off points) to ascertain the best value to ensure that meaningful pause data was captured. For the speaker in this study, the researcher found 300 ms to be the best value to establish as a cut-off point. A duration of 400 ms eliminated many of the normal sentence juncture and pauses before clauses. Likewise, values below 300 ms were too sensitive (including 250 ms), tremendously increasing the number of pauses detected, even in positions inconsistent with our knowledge of normal pausing (see Chambers, 1997). Having chosen the cut-off point of 300 ms, further analysis was carried out to investigate whether the speaker's pausing conformed to standard pausing patterns as defined by Chambers (1997). Pauses were categorized by types including sentence juncture, before difficult lexical items, before relative clauses, before prepositional phrases, and other phrase and clause boundaries (see Table 1).

Table 1

Passage Pause Analysis Breakdown by Type and Mean Duration

Location	N	Percent	Mean Duration
Clausal	17	30.36%	0.641499
Lexical	15	26.79%	0.656934
Preposition	4	7.14%	0.4515
Sentence	20	35.71%	0.79906

Pauses were distributed among these four common positions. As expected, pauses that occurred at sentence boundaries had the longest average duration. To determine which cut off point was

appropriate, the passage text was checked for pausing that should occur, (i.e. sentence juncture) but that were not captured by one of the particular cut-off points chosen.

Having chosen a cut-off point of 300 ms, the mean for each pause type was calculated. To create different versions of the passage, the researcher choose to modify pause duration by a percentage rather than an arbitrary value, allowing for the pauses to preserve their natural characteristics. Pauses at sentence boundaries are generally higher in duration than pauses at clausal boundaries, so the researcher wanted to modify pause duration in a manner that preserved differences such as this. The value of 50% was chosen for both shortening and lengthening the pause lengths in the unmodified passage. This percentage of change was used in order to maintain the natural pause locations of the speaker while falling within pause durations suggested by the research (see Riggenbach, 1991). The researcher made careful notes concerning the location of each pause by noting the initial and final words of each thought group. Each pause was then manipulated using Audacity[1], an open-source audio editing tool, to the designated reduced or lengthened values. To lengthen the pause duration values, a period of silence was added to the original pause to reach the desired value. In order to shorten pauses, the difference between the original duration and the desired shortened value was deleted from the center of the original pause. The researcher then listened to each passage to verify that the passages were modified correctly. Table 2 shows the differences between the three versions of the passage.

Table 2

Differences in Pausing by Passage

	Mean Pause Length	Min Pause Length	Max Pause Length	Total Pause Duration	Length of Passage	Percent of passage in pauses
Short	0.263527	0.15	0.58	14.49 s	3:05	7.83%
Medium	0.527054	0.30	1.16	28.99 s	3:18	14.64%
Long	0.790582	0.45	1.74	43.48 s	3:33	20.41%

Note: Each passage variation contained a total of 55 pauses.

The normal range of pauses included pauses from a minimum duration of 0.3 seconds to a maximum of 1.16 seconds, with a mean pause duration of 0.53 seconds. Additionally, manipulating the pauses caused the length of the passage to range from 3:05 to 3:33 in duration.

Survey. Once the different versions of the passage were created, items were created to gauge intelligibility and evaluative reactions towards the speaker. A comprehension quiz consisting of 10 questions was created to measure listener's understanding of the main points in the lecture. The questions were modeled after the comprehension questions used in Hahn's 1999 dissertation work. Questions were then evaluated by a testing professional and a class of TESOL masters graduate students enrolled in a testing course. Questions were modified, altered, or deleted based on their compliance with best listening comprehension testing practices. One key practice observed in the creation of the questions was to avoid testing text close to the beginning of the passage. Freedle and Kostin (1996, 1999) examined 337 TOEFL multiple choice items to determine what made items difficult. The result of the study was to create an inventory of characteristics that made items more difficult. Among these items, the most important contributing factor was the location of the necessary information. In accordance with literature asserting that information at the beginning of passage is easier to remember (Buck, 2001, p. 153), questions addressing the first portion of the passage were removed. Additionally, the list of

questions were presented to a panel of three linguists, critiqued, altered and improved. The final

list of questions consisted of six questions – each focused on different sections of the listening

passage (see Table 3).

Table 3

Comprehension Questions

Number	Question
1	Why does the speaker use the example of ripening bananas?
2	According to the speaker, what is the role of carbohydrates?
3	Why is protein essential for our body?
4	Why does the speaker say that lipids are the "notorious nutrient"?
5	In what ways are lipids essential for sustaining our lives?
6	Why do food scientists need to understand the three basic nutrients?

All six questions focus on main ideas presented in the passage. Effort was made to ensure that

raters answered the passage based on information from the passage, thus a number of questions

refer directly to the lecture.

Next, an open-ended recall item was also included to investigate intelligibility by

measuring rater recall of main ideas from the passage. The prompt was as follows: "Write down

everything you can remember from the lecture. You will have three minutes. Write for the whole

time period. Include specific details that you can remember from the lecture." This open-ended

recall item was designed so raters would be given three minutes as a time limit to respond to the

prompt, encouraging them to provide complete and full responses. Raters were not able to

advance until the three minutes had completely expired, at which time the survey automatically

advanced. This open ended recall item was designed to measure the quantity of main ideas raters

were able to identify from the passage in their responses to the open-ended recall response item. This was one of the ways whereby the survey collected data concerning the intelligibility of the speaker. A total of thirteen main ideas were identified in the passage. The main ideas were identified independently from the passage by two raters with experience in linguistic analysis. Each rater analyzed a transcript of the audio passage by using methods modeled after Hahn's 1999 dissertation, which employed Johnson's Protocol (1971) to identify main ideas and supporting details. After identifying the main ideas independently, one rater had identified 12 main ideas, the other 13 main ideas. The two lists were the same except that one rater had taken one main idea and separated it into two ideas. Upon close review of the items in question, the two raters reached consensus that the passage included 13 main ideas (see Table 4).

Table 4

Main Ideas Identified from the Passage

Number	Main Idea
1	Food science explains how foods can be kept fresh in the winter
2	Food science improves preservation of food
3	Nutritional content is improved through food science
4	Flavor is improved through food science
5	Gas is used to artificially ripen different fruits and vegetables
6	There are three main nutrients: Carbohydrates, lipids, and proteins
7	The unique qualities of carbohydrates
8	Proteins are used by the body for growth
9	Proteins are unique for containing nitrogen
10	Why lipids are considered negatively
11	The ways lipids have a positive influence on our body
12	Lipids help to store energy
13	The three main nutrients interact with one another

In order to remain consistent with the literature (see Buck, 2001, p. 153), the first main idea was eliminated from comprehension measures using the notion that the first idea of a passage is the most easily recalled by listeners, resulting in a total of 12 main ideas for the passage.

The next part of the instrument included a series of items designed to elicit raters'
evaluative reactions of the speaker. All questions were modeled after questions asked in Hahn's
(1999) dissertation. These items consisted of nine binary choice items and three open-ended
items. Finally, items were created to garner listener demographical data. (Refer to Appendix B
for the full text of the survey.) Raters' reactions to the passage and speaker were obtained
through three different question types. All questions were modeled on Hahn's 2004 work which
employed a battery of binary questions, Likert scale questions, and an open ended item (Hahn,
2004, p. 208). The binary questions are displayed in Table 5.

Table 5

Binary Evaluative Questions

Label[1]	Question
B1	The speaker's voice was 1) pleasant to listen to 2) irritating to listen to
B2	Would you like to know more about this topic?
B3	It was easy to hear and understand the instructor
B4	The instructor's lecture seemed to ramble
B5	The instructor's presentation would allow for easy notetaking
B6	The instructor's voice was pleasant to listen to
B7	The instructor's voice was irritating
B8	The instructor's English ability hindered communications of ideas
B9	The instructor emphasized important events by raising voice, repeating, etc.

Note. [1]Label references the designation of each question on the survey, see Appendix B.
The available responses to question B1 are given with the question. The available responses to
questions B2-B9 were all "yes" and "no".

The binary choice questions, labeled with B, focused on the speaker's vocal quality and English-
speaking ability (B1, B6, B7, B8, B9), the delivery of the lecture content (B4, B8, B9), and raters'
perception of ease of understanding the lecture (B3, B5), and interest in the lecture (B2).

Open ended items (see Table 6) were included to provide access to responses not considered in the creation of the selection items and provide additional rich data.

Table 6

Open-Ended Evaluative Questions

Label	Question
E1	Was the teacher a good speaker? Why or why not? (give specific reasons)
E4	Describe the instructor's ability to explain the lecture content.

The open-ended questions focused on listener judgments as to whether the speaker was competent or not, and the reasons for this (E1), as well as raters' description of the speaker's language ability (E4).

Two other Likert scale questions were included. The first, "Understanding the speaker was…" asked raters to choose on a five point level of difficulty to understand scale ranging from "very easy" to "very difficult." The second question was "How interesting was the speaker's presentation?" Raters choose from a scale of 6 points ranging from "very interesting" to "very boring."

In total, the survey consisted of 27 separate items, including a series of 9 questions in the final portion of the survey which were designed to gather information about the raters' academic backgrounds, knowledge of other languages, amount of interaction with non-native English speakers (NNSs), and their level of familiarity with the topic (see Appendix B - Task 4).

Procedure

Raters came to a controlled lab setting to complete the listening tasks. Surveys were administered through Qualtrics, an online survey creation instrument, and were taken via the Google Chrome internet browser. At log in, raters were randomly assigned by the Qualtrics

software to listen to one of the three versions of the listening passage and the accompanying comprehension questions. Next, each rater indicated his or her informed consent to participate in the study. Following an indication of consent, raters listened to a brief audio recording designed to test their hardware, after which they indicated the degree to which they were able to clearly hear the audio through their headphones. Raters all used identical headphones for the study. If raters indicated that they were unable to hear the audio, or the audio was not clear, they were not able to progress through the survey. Raters were told to notify a lab assistant if they were unable to hear the audio clearly. A second listening check was performed prior to raters listening to the main audio passage. After completing the audio test, all raters listened to a 41 second speech sample of a non-native English speaker talking about brain science. They were then asked to answer a number of items indicating their attitude towards the speaker (see Appendix B – Task 3). This passage was included in order to provide a reference point for the raters who would later rate different passages. If raters were harsh towards non-native speakers (particularly Japanese speakers) useful data would be obtained to conduct an analysis of covariance and compensate for rater differences. Following this brief exercise, raters listened to the version of the main audio passage assigned to them. They were only able to listen to the audio passage one time, then they advanced to the next page of the survey. The construct of overall comprehension was measured through two different means. First, raters completed an open-ended recall response item within a three-minute time limit. Raters were given 20 seconds to read the question and three minutes to type their response. The survey automatically advanced to the next question when the time expired. Next, raters completed the six open-ended response items to measure their comprehension of the lecture. These items were designed as measures of the intelligibility of the speaker, (if the raters were able to comprehend the content of the passage).

Raters then answered a series of choice items (see Appendix B – Task 3, items E3, E5, and B1-B9) identical to the items they completed in response to the 41 second passage at the beginning of the survey. Finally, they completed a series of questions about their experiences with non-native speakers and their academic career (see Appendix B – Task 4). Raters also supplied information about their familiarity with the topic of food science by answering a simple multiple choice question that asked "How much did you know about the topic of food science and its basic nutrients?" Raters responded by choosing from the options "well acquainted", "heard about it before", or "totally unfamiliar" (see Appendix B – Task 4).

Scoring of responses. The qualitative questions in the survey were coded for analysis. First, the open-ended recall item, the first measure of intelligibility, was coded in the following manner: Raters' responses were assigned a maximum value of 2 points each per main idea identified. The scoring system was as follows: 0 points for a completely incorrect or missing response, 1 point for partially identifying the main idea, and 2 points for completely capturing the main idea. Raters' responses were scored independently by two graduate students with backgrounds in TESOL and linguistics, yielding a high correlation ($r = .89$).

The second measure of intelligibility, the comprehension quiz, was also coded on a 2-point scale, with 0 given for a response that was completely incorrect, 1 for a partially correct response, and 2 for a completely correct response (see Appendix C). Inter-rater reliability was calculated on the scoring with a high correlation ($r = .87$).

Two evaluative questions about the speaker also provided qualitative responses. These questions were prepared for analysis by first closely examining the content of the rater responses. Specific categories of responses (for example, "good pronunciation") emerged from rater responses as the data were examined. Rater responses were coded into a spreadsheet containing

these categories using a value of 1 to mark if a category was mentioned in the rater's response and a value of 0 to indicate the absence of a category in the response. Two graduate students from the Department of Linguistics and English Language completed the inter-rater reliability of the coded data. Inter-rater reliability for these items was calculated by using a Pearson Correlation *at r =*.89.

Chapter 4

Results

This study was designed to gather data on the effect of pause duration on intelligibility and evaluative reactions to the speaker. Raters (n =95) were divided into three random groups where they listened to one of three versions of the same passage, each passage differing only with regards to pause duration (shortened, unmodified, and lengthened). Raters then completed an open-ended recall task and comprehension quiz to assess intelligibility. Next, raters provided responses to questions eliciting evaluative reactions of the speaker. Finally, raters completed a series of demographic questions to enable the researchers to identify possible extraneous or intervening variables and/or group differences.

This chapter will present the results of the study by first examining the analysis of differences between groups with the demographic data collected. Following this, the chapter will provide the results of the analysis of the recall task and comprehension quiz. Finally, analysis of the evaluative items will be presented.

Analysis of Differences between Groups

To ascertain whether there were significant differences between the three groups that acted as confounding variables, a Chi Square analysis was completed for each of the items that collected demographic data about the raters. Chi Square yielded no significance between groups for the items interaction with non-native speakers ($\chi^2(2) = 2.93$, $p = .23$), experience living in another country ($\chi^2(2) = 0.003$, $p = .99$), and additional languages spoken ($\chi^2(2) = 1.33$, $p = .51$). Raters were also found not to differ significantly with regards to their tolerance of non-native English speaker's language [$F(2,92) = 0.56$, Pr>F = .57]. Familiarity with the topic was also assessed and found to have no statistical significance (see Table 7).

Table 7

Familiarity with Topic

Version	N	Mean	SD	Pr > |t|
Long (A)	32	1.97	0.0098	<.0001
Normal (B)	32	2.22	0.0098	<.0001
Short (C)	31	2.06	0.0099	<.0001

Note: All groups did not differ significantly in any of the areas measured.

Familiarity with the topic is an issue of great concern to the validity of the study. However,

between groups, the mean values are not measurably different enough to impact the results of the

study.

Measures of Intelligibility

Recall task. A one-factor ANOVA of the main ideas recalled by the raters showed no

significant difference between the three groups [$F(2,92) = 0.52$, Pr>F = .60]. A Tukey-Kramer

posthoc analysis was applied to the data. The Tukey-Kramer analysis (also known as Tukey's

Honestly Significant Difference Test) compares all possible pairs of means, and is based on a

partipantized range distribution. Effect size was also calculated to show the differences between

the groups without confounding the difference with sample size (see Table 8).

Table 8

Recall Task

Version	N	Mean	SD	Effect size
Long (A)	32	6.81	0.68	1.43 (A-B)
Normal (B)	32	7.78	0.68	0.78 (B-C)
Short (C)	31	7.26	0.69	0.66 (A-C)

Despite the absence of statistical significance, the mean scores for each passage follow a trend

that suggests a practical difference. The mean value for the normal passage is almost a full point

(.97) higher than the passage with lengthened pauses. Raters of the normal passage also

performed better at recalling main ideas than those who listened to the passage with shortened

pauses. It is quite possible that an increase in the number of raters could yield significant results.

Analysis of the recall of each main idea was also carried out to ascertain the features that

increased or decreased comprehension of the main ideas in the passage. Of all the main ideas

identified from the passage (M1-M12), the question begs to be asked which were the most salient

and why? Table 9 presents data concerning the main ideas analyzed, the number and nature of

rater responses, and data about pausing within each main idea group.

Table 9

Number of Main Ideas Recalled with Pause Frequency and Duration

Main Idea	Key Words	Incorrrect (0)	Partially identified (1)	Identified (2)	% of Raters Identified	Pause Frequency	Total Pause Duration
M1	Preserved	74	3	18	22%	2	0.952
M2	Nutrition	71	2	22	25%	3	1.408
M3	Flavor	76	0	19	20%	1	0.432
M4	Gas to Ripen	14	10	71	85%	9	4.738
M5	Three Nutrients	25	17	53	74%	4	2.918
M6	Carbs/Hydrogen	84	0	11	12%	3	2.634
M7	Protein Growth	73	0	22	23%	4	3.76
M8	Nitrogen	78	2	15	18%	2	1.84
M9	Lipids Negative	43	16	36	55%	4	3.08
M10	Lipids Positive	59	6	30	38%	4	2.744
M11	Store Energy	85	0	10	11%	2	1.736
M12	Interaction	84	0	11	12%	5	3.976

The data indicate that M4 and M5 had the highest percentages of correctly identified main ideas

and partially identified main ideas (85.76% and 73.68%, respectively). A Pearson Correlation

was calculated for each correctly identified main idea and the total duration of pauses within the

main idea group. A significant value of $r = .53$ was identified. This suggests that longer pause

duration occurring within a main idea appears to have contributed to an improved recall of that main idea.

Comprehension quiz. Performance on the comprehension quiz analyzed by a one-factor ANOVA did not yield statistically significant results [$F_{(2,92)}$ = 2.07, Pr>F = .13].

Analysis of Evaluative Data

Binary-choice items. Evaluative data obtained through the battery of nine binary-selection items (see Appendix B – Task 3 or Table 5 for a complete list of questions) were split into two categories: positive impressions of the speaker and negative impressions of the speaker. Positive impressions included six items assessing vocal quality, comprehension of the speaker's language, interest in the topic, desire to learn more, hearing the speaker, ease of understanding the speaker, ability to take good notes, pleasant vocal quality, and emphasis of important information. Negative impressions consisted of three items: Rambling, ineffective communication of ideas, and irritating vocal quality.

Positive evaluative data selection items were run through a one-factor ANOVA. [$F_{(2,92)}$ = 1.60, Pr>F = .21]. A Tukey-Kramer post hoc analysis was applied to the data. It was also employed in Hahn's 2004 study. Effect size was calculated to illustrate the differences between the three groups (see Table 10).

Table 10

Positive Impressions of the Speaker

Version	N	Mean	SD	Effect size
Long (A)	32	5.15	0.28	1.25(A-B)
Normal (B)	32	5.50	0.28	2.62 (B-C)
Short (C)	31	4.74	0.30	1.41 (A-C)

Note: There was a possible total of eight points

The data, though not significant, exhibit the same trend observed on the recall task. Mean scores for the normal, unmodified passage were higher than either the passage with shortened or lengthened pausing.

Negative impressions of the speaker were analyzed in the same manner as positive impressions, except that the values were reversed. The results for this feedback were also not significant [$F(2,92) = 0.27$, $Pr>F = .77$] (see Table 11).

Table 11

Negative Impressions of the Speaker

Version	N	Mean	SD	Effect size
Long (A)	32	.69	0.18	0.67(A-B)
Normal (B)	32	.81	0.18	0.33 (B-C)
Short (C)	31	.87	0.18	1.00 (A-C)

In this case, a mean value closest to zero indicates that the speaker was viewed more negatively. Raters of all three versions of the passage fall within a positive range. Though the differences between the mean scores are minimal, raters who listened to either the short or normal passage gave the speaker higher ratings than raters who listened to the passage containing lengthened pauses.

Qualitative items. Responses to two open ended questions, "Was the teacher a good speaker? Why or why not? (give specific reasons)" and "Describe the instructor's ability to explain the lecture content" were analyzed first using the same measures as Hahn used in her study. A multivariate ANOVA was performed with 14 categories of responses as the dependent variables and the three versions of the passage as the factor variables. This analysis was conducted to determine whether differences were present in the three groups' responses to the multiple dependent variables. Results indicate a significant difference for the various conditions

[F(2,92) = 1.59, p = .04]. The MANOVA shows that differences are present between the three groups. The p-value is relatively high for overall difference. Due to the type of data collected, Chi Square was deemed a better measure of analyzing the differences between groups based on their responses to both of the qualitative prompts. Chi Square was chosen because the data have two groupings associated with each other: First, the presence or absence of a particular item in the qualitative response, and second, the variation of the passage to which raters listened. Of the 14 categories, two were found to be statistically significant, and the trend in an additional category, intonation, is suggestive. Categories which did not achieve significance did not have a sufficiently high frequency of response. Results of the Chi Square are found below in Table 12.

Table 12

Chi Square Analysis of Qualitative Items by Passage

Item #	Item name	Df	Value	Short[1]	Normal[1]	Long[1]	Sig.
1	Good speaker	2	2.603	22	28	25	0.271
2	Pronunciation	2	9.613	4	13	4	0.008*
3	Ease of understanding	2	3.903	3	2	7	0.142
4	Speech rate	2	3.476	2	1	5	0.176
5	Clarity	2	4.339	6	10	14	0.114
6	Intonation	2	6.410	8	6	1	0.040*
7	Good organization	2	1.317	5	7	9	0.518
8	Monotone	2	1.155	3	4	6	0.561
9	Boring	2	0.231	4	3	4	0.890
10	Bad organization	2	3.0268	5	4	1	0.220
11	Overall good impression	2	10.145	15	25	13	0.006*
12	Clear and concise	2	1.737	4	8	5	0.419
13	Knowledgeable	2	0.4973	7	6	5	0.779
14	Lack of interest	2	0.505	3	4	5	0.777

Note: * indicates statistical significance.
    [1] indicates the frequency of responses by raters to the passage type to which they listened.

Items two, six, and eleven reached statistical significance. It should be noted that the p value of item six is borderline significant at the .05 level based on the number of responses present. For the three response categories, significant group differences are displayed in Table 13.

Table 13

Pairwise Results for Significant Qualitative Items

Item #	Item name	p value	Significant group differences
2	Pronunciation	0.008	Normal>Long, Normal>Short
6	Intonation	0.040	Normal>Long, Short>Long
11	Overall Good Impression	0.006	Normal>Long, Normal>Short

Raters of the normal passage commented more frequently in a positive manner about the speaker's pronunciation than both the passages with shortened and lengthened pauses. Raters who listened to the normal passage commented positively 13 times (41%), while those who listened to the lengthened passage and shortened passage each only commented 4 times (13%). Positive comments about the speaker's intonation were most frequent in the normal and surprisingly, the shortened passage. However, it should be noted that a p value of less than 0.01 would be needed to establish strong statistical significance based on the number of responses. Raters who listened to the normal, unmodified passage gave more positive comments (item number 11) than either the passage with lengthened or shortened pausing. Raters who listened to the unmodified passage gave 25 (of a possible 32) positive responses (78%), to the lengthened passage there were 13 (out of a possible 32) (41%), and to the shortened passage, a total of 15 (out of a possible 31) (47%).

Chapter 5

Conclusion

The purpose of this study was to investigate the effect of pause duration in spontaneous oral speech on the raters' intelligibility of the passage. The study also explored pause duration's impact on evaluative reactions to the speaker by the raters. Analysis of intelligibility and rater's evaluative reactions were performed through the use of an ANOVA and Chi-Square. A Pearson Correlation was also calculated for the number of raters who successfully recalled main ideas in conjunction with the passage's rate of pause duration.

Analysis of an open ended recall task showed that raters who listened to the normal, unmodified passage, were slightly better able to recall main ideas than their counterparts who listened to the other two modified versions of the passage. However, the difference did not reach levels of statistical significance. Scores on the comprehension quiz followed the same trend, but were also not significant (p = .61). A positive correlation (r =.53) was found between the number of raters able to recall main ideas and the total pause duration within the main idea group. This finding suggests that total pause duration appears to play a role in making main ideas in a passage more salient to a listener.

Evaluative reactions to the speaker were measured in two ways: 1) Through responses to binary and Likert scale items and 2) Open-ended qualitative responses. An ANOVA showed that differences between the three groups were not significant with regards to the binary and Likert scale questions. Significant differences were found within the qualitative comments. Specifically, both positive reaction to the speaker (p = 0.006), and overall impression of pronunciation (p = 0.008) were measurably different between the three groups. Positive reactions to the speaker were measurably more frequent in the passage containing normal pausing as

opposed to either the passage with shortened or the passage with lengthened pauses. Rater's comments about overall pronunciation followed the same trend, being more frequent in qualitative responses to the passage containing normal, unmodified pauses than in either of the other two versions of the passage.

This chapter will proceed by first presenting the implications of these findings and discussing their relation to the research questions of this study, which are as follows: (1) What effect does the duration of pausing at sentence and clausal boundaries within spontaneous oral discourse have on the intelligibility of non-native speech? (2) Is there optimal pause duration for maximum intelligibility in non-native speech? (3) What effect does pause duration have on evaluative reactions toward the speaker of the passage? This discussion is followed by a discussion of the limitations in this study. The chapter concludes with suggestions for further research.

Implications

Intelligibility. The first research question asked in this study was, "What effect does the duration of pausing at sentence and clausal boundaries within spontaneous oral discourse have on the intelligibility of non-native speech?" This study's focus on spontaneous oral discourse allows for implications that have not been drawn in prior research.

The variable of spontaneous oral discourse is essential for understanding the results of the study. Studies such as Ruder and Jensen (1972), Duez (1982), Oliveira (2002) have investigated the nature of pausing in spontaneous speech. Other researchers (see Raupach, 1987 and Riggenbach, 1991) have focused on the connection between fluency and pausing, but few have examined the effects of pause duration in spontaneous oral discourse. Derwing et al. (2004) is one of the few studies to study intelligibility and pauses, establishing a connection between

pause frequency and the intelligibility of the speaker (p. 669). The current study provides

empirical data to confirm that there is a connection between pause duration and intelligibility

ratings. The trend in the data gathered by the open-ended recall item aligns with that found in

the Derwing et al. (2004) study. It is helpful to examine the differences between the three

passages to ascertain the reason why scores on intelligibility items vary. Table 2 is reproduced

below to provide detailed information concerning the differences between the three versions of

the passage.

Table 2

Differences in Pausing by Passage

	Mean Pause Length	Min Pause Length	Max Pause Length	Total Pause Duration	Length of Passage	Percent of passage in pauses
Short	0.263527	0.15	0.58	14.49 s	3:05	7.83%
Medium	0.527054	0.30	1.16	28.99 s	3:18	14.64%
Long	0.790582	0.45	1.74	43.48 s	3:33	20.41%

Note: Each passage variation contained a frequency of 55 pauses.

The normal range of pauses included pauses from minimum duration of .3 seconds to a

maximum of 1.16 seconds, mean pause duration of .53 seconds. What does this tell us about the

normal, unmodified pauses and their relationship to the pause modified by 50% in the other

versions of the passage?

Fayer and Krasinsky's 1995 exploration of intelligibility and pausing ranked seven

speakers by intelligibility ratings. They found that the data suggested that high intelligibility

rankings exhibited a strong negative correlation with percent pause time, length of the average

pause, and duration of the longest pause. Comparing the three passages used in this study yields

interesting results. In Fayer and Krasinski's study, the highest ranked speaker spent 8% of his or

her time in pauses, had an average pause length of 0.95 seconds, and the duration of the longest

pause was 1.26 seconds (p. 117). In the shortened version of the passage of the current study, the same percentage of time spent in pauses was also 8%, however, the average pause length and duration of the passage (with shortened pauses) were much lower (.26 seconds, .58 seconds). Based on these results, the trend in Fayer and Krasinski's data would suggest higher comprehensibility ratings for the short passage. In the analysis of negative impressions of the speaker, the shortened version of the passage did garner the most positive responses, but did not approach significance. However, this trend of higher ratings for the shortened version of the passage did not extend to all of the data. Indeed, in all other measures of comprehensibility, (positive impressions of the speaker and qualitative responses) the data indicate that raters preferred to listen to the normal and lengthened versions of the passage, rather than the passage containing shortened pausing. There may be a point when pause duration becomes too short and negatively impacts comprehensibility, but with the data of the current study, it is impossible to decisively confirm or reject this idea.

Optimal pausing. Learning more about what constitutes optimal pausing was another goal of the study. Owing to the fact that the three versions differed only with regards to pause duration, it is possible to learn more about optimal pausing. The positive correlation between length of pause duration and identification of passage main ideas also confirms and supports research on the importance of thought groups delineated by the correct pausing patterns (see; Andrade, 2005; Gilbert, 1993; Prator, 1985). This finding also seems to suggest the presence of optimal pause duration. While this finding can be somewhat misleading by suggesting that the longer the pause the greater the ability of the interlocutor to identify main ideas, it should be noted that the length of pauses examined in this study, even at their minimum and maximum, did not exceed values that could be considered as natural or acceptable. In fact, maximum pause

values of the lengthened passage in this study fall between the intelligibility values achieved by the two highest ranked speakers by Fayer and Krasinsky (1995, p. 117). This means that based on the pause durations of all three versions of the passage, the speaker would be considered a highly fluent speaker of English.

Responses to qualitative questions regarding the ability of the speaker and presentation of the lecture revealed however, a higher number of responses related to irritation and boredom in the version of the passage with lengthened pauses than in the other two passages. One rater remarked, "He (the speaker) seemed to be able to explain things well, but I had a hard time focusing because he was somewhat monotone." Perhaps a further increase and reduction of pause duration would have yielded more like-minded comments and ratings. Effort was made, however, to ensure that even modified pauses were within range natural in spontaneous speech. Indeed, according to Marcias (2006), pauses ranging from .25 to 2 seconds are natural. With the exception of 14 pauses in the short passage, the remaining 151 pauses fell within this range. It may be difficult to identify significance within what can be considered natural pausing. Blau (1990) and Reid (2003) operated studies with pauses increased by 300 ms in oral reading passages. Blau compared comprehension scores of students who listened to three variations of the same passage, one version of the passage unmodified, the second modified to contain 300 ms pauses, and the third version being mechanically slowed down. Participants in the group that listened to the passage containing 300 ms pauses performed better on a comprehension quiz than their counterparts who listened to the other two versions of the passage. Reid's study (2003) also employed 300 ms pauses, but was not able to establish significance with regards to comprehension (intelligibility). However, Reid found that pauses helped raters recognize organizational transitions within the listening passage, such as when the speaker moved between

different topics. Perhaps by increasing the modification percentage in spontaneous speech, it would be possible to demonstrate a greater connection between pause duration and intelligibility. It could be that listeners are more tolerant up to a point of lengthy pause duration in spontaneous speech than in oral readings.

The connection between pausing and intelligibility as succinctly stated by Chen (1996) is interesting in the context of this study. Chen states that when pausing is not sufficiently implemented "the listener will have difficulty processing the meaning of the sentences, even if the individual words are pronounced well" (p. 12). Though statistical significance was not shown with regard to intelligibility scores, qualitative comments seem to suggest that this is the case. One rater who listened to the shortened passage remarked that "the lecture...rambled a little bit from topic to topic more quickly than would be useful for really absorbing the information." This comment suggests that pause time enables the listener to absorb or process the content of the speech stream. Ruder and Jensen's (1972) finding that acceptable pause duration increased after using more complex syntactical structures (p. 49), concurs with this sentiment. The more complex the syntactic structure, the more imperative appropriate pausing becomes. De Pijper (1994) supports this claim as well by commenting that linguistic transparency is enhanced though appropriate pause duration and frequency (p. 2037). Using pauses appropriately allows the speaker to convey ideas to the listener with a minimum loss of information in the process.

Another way to examine the difference between the three passages and the search for optimal pausing is by applying Riggenbach's (1991) hierarchy of hesitation phenomena (p. 426). Table 14 contains the hierarchy as presented in Riggenbach's (1991) article.

Table 14

Riggenbach's Hierarchy of Hesitation Phenomena

Name	Duration
micropause	< 200 ms
hesitation	200-500 ms
unfilled pause	>500 ms
filled pause	varies

Note: filled pauses can be further classified into non-lexical, sound stretches and lexical types.

Due to the fact that pause duration, in this study, was elongated and reduced by 50% of the normal duration, according to Riggenbach's classification reported in Table 14, the pause durations used in this study would fall within different categories. In the version of the passage containing shortened pauses, Riggenbach would classify the mean pause duration as hesitations. The mean values of the normal and lengthened versions would both fall within the unfilled category, though the normal version also includes a high frequency of hesitations (see Table 15).

Table 15

Pause Duration Characteristics by Passage

	Short	Classification	Normal	Classification	Long	Classification
Mean	0.263527	hesitation	0.527054	unfilled	0.790582	unfilled
Min	0.15	micropause	0.3	hesitation	0.45	unfilled
Max	0.58	unfilled	1.16	unfilled	1.74	unfilled

By Riggenbach's classification, pauses in this study would fall into drastically different categories depending on the version of the passage. The frequency of each type of pause by passage is shown in Table 16.

Table 16

Number of Each Type of Pause by Passage Type

	micro	hesitation	unfilled
Short	14	40	1
Normal	0	31	14
Long	0	6	49

This analysis provides a new look at the three versions of the passage. If indeed, the

nature of a pause changes depending on its duration, as suggested by Riggenbach, then different

classifications of pausing would be employed to different effects in the speech stream. The

differences evident from the data of this study, chiefly comprehensibility ratings and raters'

ability to identify main ideas from the passage, seem to support Riggenbach's classification.

Raters were better able to identify main ideas if they listened to the passage with normal,

unmodified pauses, which contained a higher ratio of hesitation pauses than the other passages.

The data gathered as a result of this study indicate that different pause types have varying effects

on comprehensibility and intelligibility scores. Riggenbach remarks concerning *micropauses* that

"it is possible that such short pauses are not indicative of a lack of fluency or of breakdowns in

speech but rather are within the range of normal or fluent speech" (p. 426). The trend in the data

of this study suggests that this is true, as long as micropauses are employed in the correct

positions in a sentence. For example, micropauses in sentence final positions would reduce the

ability of the interlocutor to uptake information and/or interpret meaning.

This study adds more somewhat to our knowledge of what constitutes an optimal pause,

but does not provide a definitive answer. Disagreement in the literature concerning the nature of

pausing highlights the difficultly of establishing an arbitrary cut-off value for pause duration.

Indeed, though we can define a suggested range, the acceptable range of pausing displays a

certain amount of elasticity to compensate for difference in idiolect. Ruder and Jensen (1972) found that considerable difference in raters made it difficult to construe an accurate definition of what it means to pause appropriately.

Evaluative reactions to the speaker. The final research question in this study was "What effect does pause duration have on raters' evaluative reactions toward the speaker of the passage?" This study provides strong data to confirm and extend the results of prior research. The data complements the research of Rossiter (2009), which indicated that pausing accounted for half of all negative temporal impressions of the speaker (p. 409). Based on this finding, the researcher in the current study expected that evaluative ratings of the speaker would vary depending on pause duration. The quantitative data derived from the selection items did not yield statistical significance, but the trend in the data seems to support Rossiter's finding. Positive impressions, though also not significant, exhibited a stronger trend than negative impressions. Indeed, within the qualitative responses, positive reaction to the speaker was measurably higher for the normal passage than for the other versions of the passage. Rater evaluative reactions to the speaker and language varied according to pause duration. The surprising variance of 15 positive responses in the shortened version of the passage compared to 25 positive responses found in the normal version indicates that listeners prefer listening to the speaker when his pauses are within the normal duration. Pause duration appears to be important for establishing mutual understanding between speaker and interlocutors. When appropriate pause duration is maintained, interlocutors find it easier to listen to the speaker.

Additionally, the significant difference between the normal and modified groups with regards to perceived pronunciation is another important finding. The unmodified passage received a much higher number of comments related to good pronunciation than either of the

other passages. Though pronunciation of segmentals and suprasegmentals remained unchanged (with the exception of pause duration), raters felt that pronunciation was better in the passage containing unmodified pauses. This finding confirms prior findings concerning the paramount importance of pausing and thought group creation. Rater perception of the speaker does appear to influence their investment into comprehending the message.

Limitations

There are three major areas of limitation in this study: issues related to the passage, survey questions, and raters. This section will discuss the limitations in each area and offer suggestions for overcoming them in future studies.

Passage. The nature of the audio passage may have been a limitation in itself. As in Hahn's study, one passage was used and it was manipulated for one prosodic feature in an effort to focus attention on the influence of pause duration. The three versions of the passage were only different by having the pause duration either increase or decrease by 50%. By trying to keep the pauses within the natural range of the speaker, the researcher may have created some limitations. If pause duration had been increased and decreased by greater values, it may have been easier for raters to observe differences between the three groups. This said, the passage may have been too long (approximately three minutes) to sufficiently hold the attention of the raters and remain in their memory for an adequate amount of time to complete the various tasks required of them. The topic of the passage may have contributed to the results obtained. While the selection of this topic was intended to be unfamiliar to the listeners in the College of Humanities, there was some indication that raters' ability to recall main ideas, such as main ideas four and five, may be tied to background knowledge. This topic was also selected to measure

listeners' ability to recall complex academic information. Different results might be obtained if the passage was less complex, such as a personal narrative.

Questions. Another limitation in the study was the nature of some of the questions asked in the survey. Though the questions were carefully prepared, edited and reviewed, and previewed in a pilot study, one item on the comprehension quiz received limited responses due to issues with background knowledge. Question 3 (see Table 3) on the comprehension quiz, which asked about the role of proteins, was one such question. In rater responses, it was clear that most raters relied on their background knowledge, providing responses not connected to the passage itself. Raters commented that protein is used after workouts for building muscles, though this was never mentioned in the audio passage itself. In the analysis, all comprehension questions were analyzed independently to counter this problem. Other questions avoided this problem by referring to the passage, or other details that would not be part of common knowledge.

An additional limitation due to the nature of the questions was apparent in the analysis stage of the study. Although modeled directly after evaluative questions used in Hahn's 2004 study, the questions requiring binary type responses used in this study did not provide data as rich as data that may have come from other types of questions. Raters responding to these questions had to choose between the polarizing responses, eliminating possibly meaningful data from the analysis. Though it increases the burden of the rater, evaluative reactions of the speaker could have been better analyzed with more Likert-style or multiple choice items containing three or more possible choices. The most meaningful data gathered in this study concerning evaluative reactions was garnered from the qualitative responses.

Raters. Another area of limitation centers on the participants in this study. First, participants all came from the College of Humanities at Brigham Young University so as to limit

their familiarity with the topic of the listening passage. However, in so doing, many of the raters came from the Department of Linguistics, and are different from the general undergraduate student populace with regards to their awareness of the English language. The raters may have been more sensitive to segmental and suprasegmental features causing them to rate the speaker differently than they otherwise may have. The raters may also have been better aware of the features of discourse that signal changes in topic and alert the listeners to the presence of important information.

Additionally, though the number of raters was sufficient to conduct the study, the low number limited the analysis and conclusions available for the study. Though many sections of the analysis showed a trend, the number of raters in each group was not sufficient enough to provide solid evidence to the research questions.

Suggestions for Future Research

In reaction to the implications and limitations of this study, several suggestions for future research arise. First, a similar study using different manipulation values would provide additional information about the nature of pauses. The speaker of the passage in this study was already at a high level of English fluency, with low pause durations. Though high values may not have been natural for him, they are for other English Language Learners. If this study had used a speaker at a lower level of English proficiency, the variation in pause duration may have had a greater effect on measures of intelligibility. Increasing the percentage by which pause duration is manipulated may increase meaningful data gathered from this type of study. Increasing the pause duration values might result in identifying the point at which pause duration significantly impacts intelligibility. Such research could also help identify when pause duration

influences the rater/listener to begin thinking negatively about the speaker. This information could then be applied to our knowledge of image and identity's roles in actual intelligibility.

Future research could also focus on whether appropriate pause duration instruction has a positive role in increasing the intelligibility of non-native speech. Using a pre- and posttest methodology, teachers could provide focused instruction on thought groups and appropriate pause duration between phrases. Post test data could then be used to investigate whether or not instruction tied to the appropriateness of pause duration had a measurable effect on listeners' evaluation of NNESs' intelligibility.

Since the majority of studies on intelligibility have focused on academic contexts, one further possibility would be to investigate the issue of pause duration using other types of genres such as stories. Though this study was intended to investigate intelligibility in an academic context, further studies could employ passages created on a similar topic, but investigate how pause duration impacts listeners' abilities to recount information based on whether the spontaneous oral discourse involves telling a story or giving a lecture.

Conclusion

This study has contributed insights into the role of pause duration in spontaneous oral discourse. The empirical data from this study provides confirmation that pause duration is not only connected to fluency, but also to the way speech is interpreted and how the speaker is viewed. The quest for enhancing NNESs' intelligibility has led practitioners engaged in communicative pronunciation teaching to emphasize prosodic features over segmental features (Levis, 2005, p. 369), meaning an increased focus on prosody rather than on each phoneme present in English. This study of a specific prosodic element, pause duration, combined with research such as Hahn (2004) involving primary stress, provides further clarification regarding

aspects of prosodic features that practitioners need to include in their attempts to improve NNSs'

intelligibility.

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529-555.

Anderson-Hseih, J., & Koehler, K. (1988). The efffect of foreign accent and Speaking rate on native speaker comprehension. *Language Learning*, *38,* 561-613.

Andrade, M. (2005). Performance of Japanese EFL learners on a test of English syllabification, word stress, sentence focus, and thought groups. *Sophia University Junior College Division Faculty Bulletin, 25*, 1-25.

Bada, E. (2006). Pausing, preceding and following 'that' in English. *ELT Journal, 60*, 125-125-132. Retrieved from http://search.proquest.com/docview/85650130?accountid=4488

Barik, H. (1968). On defining juncture pauses: A note on Boomer's "Hesitation and grammatical encoding". *Language and Speech, 11*, 156-159.

Blau, E. (1990). The effect of syntax, speed, and pauses on listening comprehension. *TESOL Quarterly, 24*, 746-753. Retrieved from

http://search.proquest.com/docview/58242542?accountid=4488

Boersma, P., & Weenink, D. (2012). Praat (Version 5.3.27) [Software]. Latest version available for download from www.praat.org.

Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

Chafe, W. (1985). Some reasons for hesitating. In D. Tannen, & M. Saville-Troike (Eds.), *Perspectives on Silence* (pp. 77-86). Norwood, NJ: Ablex.

Chambers, F. (1997). What do we mean by fluency? *System, 25*, 535-544.

Crystal, D. (1987). *The Cambridge Encyclopedia of Language.* Cambridge: Cambridge
University Press.

Chen, C. (1996). *A new perspective on teaching English pronunciation: Rhythm.* Retrieved from
http://search.proquest.com/docview/62676816?accountid=4488

Chen, H. (2010). Second language timing patterns and their effects on native listeners'
perceptions. *Concentric: Studies in Linguistics*, *36*, 183-212.

Derwing, T., Munro, M., Thomson, R., & Rossiter, M. (2009). The relationship between L1
fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*, 533-557.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching:
A research-based approach. *TESOL Quarterly*, *39*, 379-397.

Derwing, T., Munro, M., & Wiebe, G. (1997). Pronunciation instruction for "fossilized"
learners: Can it help? *Applied Language Learning*, *8*, 217-235.

Derwing, T., & Rossiter, M. (2003). The effects of pronunciation instruction on the accuracy,
fluency and complexity of L2 accented speech. *Applied Language Learning, 13*, 1-18.

Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency:
Judgments on different tasks. *Language Learning*, *54*, 655-679.

Duez, D. (1982). Silent and non-silent pauses in three speech styles. *Language and Speech*,
25(1), 11-28.

Fayer, J., & Krasinski, E. (1995). Perception of hesitation in non-native speech. *Bilingual
Review, 20*(2), 114-120.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, *39*,
399-423.

Fujisaki, H. (2004). Information, prosody, and modeling – with emphasis on tonal features of speech. *Speech Prosody 2004*, (p. 1-10). Nara, Japan.

Gilbert, J. (2005). *Clear Speech: Pronunciation and listening comprehension in North American English: Student's book* (3rd ed.). New York: Cambridge University Press.

Gilbert, J. (1993). *Clear Speech: Pronunciation and Listening Comprehension in North American English, Teacher's Resource Book*. Melbourne: Cambridge University Press.

Gilbert, J. (1983). Pronunciation and listening comprehension. *Cross Currents, 10*, 53-61. Retrieved from http://search.proquest.com/docview/58133047?accountid=4488

Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech, 4,*18-26.

Griffiths, R. (1991). Pausological research in an L2 context: A rationale, and review of selected studies. *Applied Linguistics*, *12*, 345-364.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*, 201-223.

Hahn, L. (1999). *Native speakers' reactions to non-native stress in English discourse.* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*, 145-164.

Krivokapic, J. (2004). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*, 162-179.

Lass, N., & Prater, C. E. (1973). A comparative study of listening rate preferences for oral reading and impromptu speaking tasks. *The Journal of Communication*, *23*, 95-102.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*,

40, 387-417.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*, 369-377.

Levis, J., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal, 12*, 13-19.

Macias, M. P. (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting*, *8*, 25-43.

Oliveira, M. (2002). Pausing strategies as means of information processing in spontaneous narratives. *Speech Prosody 2002*, (pp. 539-542). Aix-en-Provence.

de Pijper, J., & Sanderman, A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America*, *96*, 2037-2047.

Prator, C., & Robinett, B. (1985). Manual of American English Pronunciation (4th Ed.). New York: Holt, Rinehart, and Winston.

Raupach, M. (1987). Procedural learning in advanced learners of a foreign language. In J. Coleman, & R. Towell (Eds.), *The Advanced Language Learner* (pp. 123-155). London: CILT.

Reid, R. (2003). The effect of digitally shortening and lengthening pauses on listening comprehension. *Dissertation Abstracts International: Section A*: The Humanities and Social Sciences, *64(*10*),*3589.

Riazantseva, A. (2001). Second language profieciency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*, 497-526.

Richards, J. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, *17*, 219-240.

Riggenbach, H. (1991). Toward an understanding of fluency: a microanalaysis of nonnative speaker conversations. *Discourse Processes, 14*, 423-441.

Rossiter, M. (2009). Perceptions of L2 fluency by native and non-native speakers of English. The *Canadian Modern Language Review*, *65*, 395-412.

Ruder, K. F., & Jensen, P. J. (1972). Fluent and hesitation pauses as a function of syntactic complexity. *Journal of Speech and Hearing Research, 15*, 49-60.

Simensen, A. (2010). Fluency: An aim in teaching and a criterion in assessment. *Acta Didactica Norge*, *4*, 1-13.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 84-119.

White, R. (1997). Back channeling, repair, pausing, and private speech. *Applied Lingusitics, 18*, 314-344.

Appendices

Appendix A - Transcription of Passage

Have you ever been to the supermarket and wondered why/1/ tomatoes or the bananas on the shelves are ripe and/2/ are fresh during the winter months? /3/ These questions can be answered through food science. /4/ Food scientists make sure that food /5/is able to be preserved for a long time/6/ and/7/ food will be able to give us /8/nutritional contents that will help us be healthy /9/and food scientists also incorporate flavor into our foods. /10/

Before I go into the basic concepts of food science, I would like to explain a few examples of /11/how food science works /12/so when we talked about tomatoes and bananas/13/ the reason why you can have those fruits and vegetables/14/ in/15/ the supermarket fresh during the winter is because /16/behind the scene we have tomatoes and bananas that are not ripe. /17/ When food scientists or when the supermarket manager feels that it is time for them to put them on the shelf,/18/ they place those items into a room /19/and they seal it, and they flush it /20/with ethylene gas./21/ By doing so, it artificially helps the fruit or vegetable to ripen./22/

Now that we've given you a few examples of how food science is being applied in the industry, /23/let us talk about a few basic concepts of food science./24/ First of all, we would like to talk about the three main /25/nutrients that is present in food /26/ and they are carbohydrates,/27/ proteins, and lipids./28/ Carbohydrates is chemical compound/29/ it is composed of carbons hydrogens and oxygens./30/ And it is the main nutrient which helps our body to be able to sustain itself. /31/ Protein is another compound in foods /32/that is essential for our body to be able to grow, /33/ to make new cells, to make new muscles, to make new /34/parts of the body that die constantly./35/ Proteins /36/are unique in sense that they are the only compounds in foods that have a nitrogen in them./37/ Lipids/38/ is another/39/ main

nutrient in food./40/ It is normally considered the notorious nutrient because it makes people fat /41/and it has a lot of correlations with diseases./42/ But fats and lipids,/43/ they are the same thing,/44/ they/45/ are very essential in sustaining our lives./46/ Lipids are things that will /47/help to lubricate our body,/48/ and also helps/49/ our body to store energy. /50/ These three components of food /51/are essential for food scientists to understand so we can know how they interact with each other./52/ As we find out they interact with each other we are able to invent new types of foods that will help to do the things that /53/ we want food to do for us, like improving nutritional content, /54/ adding better flavors /55/and preserving the food.

Appendix B **- Survey**

Q1   Informed Consent

Introduction   This study attempts to collect information about how well teachers whose native language is not English are understood in university classes.

Procedures   First you will listen to short speech sample and rate the speaker's ability. You will then listen to a short lecture (about 3 minutes). After listening to the lecture you will be asked to write down what you remember from the lecture, take a quiz on the lecture, and rate the speaker on his speaking ability.

Risks/Discomforts   Risks are minimal for involvement in this study.

Benefits   There are no direct benefits for participants. However, it is hoped that through your participation, researchers will learn more about the qualities that make speech understandable.

Confidentiality   All data obtained from participants will be kept confidential and will only be reported in an aggregate format (by reporting only combined results and never reporting individual ones). All questionnaires will be concealed, and no one other than then primary investigator and assistant researches listed below will have access to them. The data collected will be stored in the HIPPA-compliant, Qualtrics-secure database until it has been deleted by the primary investigator.

Compensation   There is no direct compensation for participating in this study.

Participation   Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to your academic status, GPA or standing with the university. If you desire to withdraw, please close your internet browser and notify the principal investigator at this email: lege.sensei@gmail.com. Or, if you prefer, inform the principal investigator as you leave.

Questions about the Research   If you have questions regarding this study, you may contact Ryan Lege, at 801-378-5818, lege.sensei@gmail.com or Mark Tanner, 801-422-8154, mark_tanner@byu.edu.

Questions about your Rights as Research Participants   If you have questions you do not feel comfortable asking the researcher, you may contact Dr. Mark Tanner, 801-422-8154. Or contact the director of Brigham Young University's Institutional Review Board. If you have any questions about your rights as a research participant you may contact the IRB Administrator at A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu; (801) 422-1461. The

IRB is a group of people who review research studies to protect the rights and welfare of research participants.

Q2 I give my consent to participate in this study.

Yes (1)

No (2)

Q3 Audio Test:

While you are completing this survey, you will be asked to listen to an audio recording. In order to confirm that you can hear the audio, click on play. Choose the appropriate choice below.

I could hear the audio clearly. (1)

I heard the audio but it wasn't clear. (2)

I didn't hear anything. (3)

Displayed if option 2 or 3 was selected:

Q4 Please make sure that your microphone is plugged in and your volume is adjusted to the appropriate level. If you need assistance notify a lab attendant.

I could hear the audio clearly (1)

I heard the audio but it wasn't clear (2)

I didn't hear anything (3)

Q39 Task 1

Listen to this short audio clip.   When you finish, click the "next >>" key and complete the attached questions.

Q47 Choose the appropriate response below:

	Very Easy (1)	(2)	(3)	(4)	Very Difficult (5)
Understanding the speaker					

was (1)					

Q41   How interesting was the speaker's presentation?

	Very Interesting (1)	Interesting (2)	Somewhat Interesting (3)	Somewhat boring (4)	Boring (5)	Very Boring (6)
The presentation was... (1)						

Q49 Mark either "yes" or "no"; in response to the statements below

	Yes (1)	No (2)
Would you like to know more about this topic? (1)		
It was easy to understand the instructor (2)		
The instructor's lecture seemed to ramble (3)		
The instructor's presentation would allow for easy notetaking (4)		
The instructor's voice was pleasant to listen to (5)		
The instructor's voice was irritating. (6)		
The instructor's English ability made it hard to understand some of the concepts he discussed. (7)		
The instructor used things like raising his voice, emphasizing important words and repeating important information to help me know what was important to listen to (8)		

One of the 3 versions of the passage is played randomly at this point:

Q5 Task 2

Click the play button to listen to the lecture. Listen carefully for the main ideas and supporting details.

Q6   Write down everything you can remember from the lecture.   You will have three minutes. Write for the whole time period. Include specific details that you can remember from the lect

Quiz

Directions: Answer the following 6 questions based on your understanding of the lecture you heard. When you are finished with the current question, click on the arrow to advance to the next. You will not be able to return to questions once you have clicked "next >>"

Q8   1.   Why does the speaker use the example of ripening bananas?

Q10 2.   According to the speaker, what is the role of carbohydrates?

Q11 3.   Why is protein essential for our body?

Q12 4.   Why does the speaker say that lipids are the "notorious nutrient"?

Q13 5.   In what ways are lipids essential for sustaining our lives?

Q14 6.   Why do food scientists need to understand the three basic nutrients?

Task 3: Evaluating the Speaker

E1 Was the teacher a good speaker? Why or why not? (give specific reasons)

B1   The speaker's voice was

pleasant to listen to (1)

irritating to listen to (2)

E3 Choose the appropriate response below:

	Very Easy (1)	(2)	(3)	(4)	Very Difficult (5)
Understanding the speaker was (1)					

E4   Describe the instructor's ability to explain the lecture content.

E5   How interesting was the speaker's presentation?

	Very Interesting (1)	Interesting (2)	Somewhat Interesting (3)	Somewhat boring (4)	Boring (5)	Very Boring (6)
The presentation was...						

Mark either "Yes" or "No" in response to the statements below

	Yes (1)	No (2)
Would you like to know more about this topic? (B2)		
It was easy to hear and understand the instructor (B3)		
The instructor's lecture seemed to ramble (B4)		
The instructor's presentation would allow for easy notetaking (B5)		
The instructor's voice was pleasant to listen to (B6)		
The instructor's voice was irritating. (B7)		
The instructor's English ability hindered communications of ideas (B8)		
The instructor emphasized important events by raising voice, repeating, etc. (B9)		

Task 4: Additional Information

Q21 Finally, we need some basic information about you.   Overall in your college experiences what percentage of your TAs and professors have been non-native speakers of English?

_____ Percentage (1)

Q22   Besides having non-native speaking instructors, have you ever had any experiences interacting on a regular basis with a non-native speaker or speakers of English?

Yes (1)

No (2)

Q23 Have you ever lived in another country or area where a language other than English is spoken?

Yes (1)

No (2)

Q24 Where? For how long?

Q47 Are you a native speaker of English?

Yes (1)

No (2)

Q25   Do you speak any language(s) other than English?

Yes (1)

No (2)

Q26 Which language(s)?

Q27   How much did you know about the topic of food science and its basic nutrients?

Well acquainted (1)

Heard about it before (2)

Totally unfamiliar (3)

Q28 Mark the appropriate response on the scale:

	Strongly Disagree (1)	Disagree (2)	Neither Agree nor Disagree (3)	Agree (4)	Strongly Agree (5)
I am patient and tolerant when interacting with non-native speakers (1)					

Q29 What is your current year in school?

1st year (1)

2nd year (2)

3rd year (3)

4th year (4)

Graduate (5)

Appendix C **- Inter-rater Reliability Guide: Coding**

Wrong (0), partial (1), Correct (2)

Note:

- On some questions, there is not an option for partial correctness(so only 0 or 2)
- Mark 2, if the response includes any of the main ideas in the suggested response
- Example: SR=makes you fat and correlates with disease, Response=correlates with disease give it a "2"

Questions

1. Why does the speaker use the example of ripening bananas?

Suggested response (2): To show how food science is applied in the world

Partially Correct (1): To keep them ripe in the winter months (or describe the process)

Not correct(0): I don't know/ I don't remember

2. According to the speaker, what is the role of carbohydrates?

Suggested response (2): helps to renew or sustain our bodies/grow new parts

Partially Correct(1): Give us energy (not according to the lecture)

Not correct(0): I don't know/ I don't remember

3. Why is protein essential for our body?

Suggested response (2): To grow, make new cells, muscles, they have nitrogen in them

Not correct(0): Gives us energy

4. Why does the speaker say that lipids are the "notorious nutrient"?

Suggested response (2): makes people fat and correlates with a lot of diseases (just one of these is okay)

Not correct: (0) I don't know any other response

5. In what ways are lipids essential for sustaining our lives?

Suggested response(2): lubricate our body, help our body to store energy

Not correct(0): I don't know/ any other response not including the above

6. Why do food scientists need to understand the three basic nutrients?

Suggested response (2): to know/understand how they interact with each other, to invent new foods, improve nutritional content, add better flavors, preserve food

Not correct(0): Any response not including one or more of the above items