Quantifying Speech Pause Durations in Typical English Speakers

John Hoffer
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Education Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/10035

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Quantifying Speech Pause Durations in Typical English Speakers

John Hoffer

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

Master of Science

Shawn L. Nissen, Chair
Connie Summers
Tyson G. Harmon

Department of Communication Disorders
Brigham Young University

Copyright © 2023 John Hoffer
All Rights Reserved
ABSTRACT

Quantifying Speech Pause Durations in Typical English Speakers

John Hoffer
Department of Communication Disorders, BYU
Master of Science

This study examines filled and unfilled pause durations between utterances in the speech of 60 people with no language disorder. It also evaluates the proportions of different pause lengths, examines the location of pauses within an isolated speech sample, and compares speech pause in male speakers and female speakers. Using speech samples gathered from a picture description task, Praat acoustic analysis software was used to segment C-units and measure pause duration between utterances (Boersma & Weenink, 2022; Öktem et al., 2021). Descriptive statistics were used to analyze these data, including pause duration mean and standard deviations. Pause mean durations ranged from 70 ms to 90 ms. Speakers used pauses shorter than 0.5 seconds and 1 second more frequently than longer pauses. Both pause frequency and mean pause length increase in the final 50% of the speech sample compared to the initial 50% for both male and female speakers. No significant differences were found between male and female speakers. Speakers produce prolongations at a rate of 0.07 to 0.08 per C-unit across both male and female speakers. Both male and female speakers have a higher frequency of prolongations in the final portion of the speech sample, compared to the initial portion. Further research across several types of speech tasks is needed to provide greater insight into variations in pause duration and location in different types of speech tasks. Further research might also examine pause durations within utterances.

Keywords: speech pause, prosody, aphasia
ACKNOWLEDGMENTS

I would like to begin by expressing my deepest gratitude to Brigham Young University for providing me with the opportunity to pursue higher education and for fostering an environment that has transformed me both personally and professionally. The knowledge, skills, and experiences gained at this esteemed institution have been invaluable in shaping my academic journey. I am immensely grateful to the Fletchers for their generous financial support. Their assistance has played a pivotal role in making my research possible, and for that, I am truly thankful.

I extend my heartfelt appreciation to my family, friends, and cohort for their unwavering support and encouragement throughout this endeavor. Their belief in me and my work has been a constant source of motivation and inspiration, and I am privileged to have such a dedicated circle of individuals by my side.

I also wish to acknowledge the tireless efforts of those who contributed to the research process. Their diligence, expertise, and commitment to advancing our collective understanding have been instrumental in the success of my thesis. I am deeply grateful for their hard work and collaboration.

I would like to extend my profound gratitude to my thesis advisor, Dr. Nissen, and the entire thesis committee for their invaluable guidance, support, and mentorship throughout this journey. Their combined expertise, constructive feedback, and encouragement have been instrumental in shaping my research and in fostering my growth as a scholar.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>DESCRIPTION OF THESIS STRUCTURE AND CONTENT</td>
<td>viii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Linguistic Factors</td>
<td>2</td>
</tr>
<tr>
<td>Perceptual Factors</td>
<td>3</td>
</tr>
<tr>
<td>Task Types</td>
<td>4</td>
</tr>
<tr>
<td>Pause Length</td>
<td>5</td>
</tr>
<tr>
<td>Speech Pause in Individuals With Aphasia</td>
<td>6</td>
</tr>
<tr>
<td>Formal Assessment of Pause in Aphasic Patients</td>
<td>7</td>
</tr>
<tr>
<td>Study Purpose</td>
<td>9</td>
</tr>
<tr>
<td>Method</td>
<td>9</td>
</tr>
<tr>
<td>Participants</td>
<td>9</td>
</tr>
<tr>
<td>Speech Task</td>
<td>10</td>
</tr>
<tr>
<td>Speech Recordings</td>
<td>10</td>
</tr>
<tr>
<td>C-Unit Segmentation</td>
<td>10</td>
</tr>
<tr>
<td>Acoustic Measurements</td>
<td>10</td>
</tr>
<tr>
<td>Reliability</td>
<td>11</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1  Identified Gender and Speech Sample Characteristics for Adult Participants .......... 26

Table 2  Silent Pause (SP) Duration and Frequency Measures Across Sample Location and Speaker Gender ......................................................... 28

Table 3  Filled Pause (FP) Duration and Frequency Measures Across Sample Location and Speaker Gender .............................................................. 29

Table 4  Prolongations Across Sample Location and Speaker Gender ........................................ 30
LIST OF FIGURES

Figure 1  Example of Speech Pause Analysis in Praat ............................................................... 31
Figure 2  Percentage of Occurrence of Silent Pauses Across Pause Length and Speech
Sample Location for Female Speakers .................................................................................. 31
Figure 3  Percentage of Occurrence of Silent Pauses Across Pause Length and Speech
Sample Location for Male Speakers .................................................................................... 32
Figure 4  Percentage of Occurrence of Filled Pauses Across Pause Length and Speech
Sample Location for Female Speakers ................................................................................ 32
Figure 5  Percentage of Occurrence of Filled Pauses Across Pause Length and Speech
Sample Location for Male Speakers .................................................................................... 33
Figure 6  Percentage of Occurrence of Prolongations Across Speech Sample Location
for Male and Female Speakers ............................................................................................. 33
DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *Quantifying Speech Durations in Typical English Speakers* is part of a larger study exploring the impact of pause on speech communication in people with aphasia. Portions of this thesis may be submitted for publication, with the thesis author being included in the list of contributing coauthors. An annotated bibliography is provided in Appendix A, and the BYU Institutional Review Boards’ approved consent form is provided in Appendix B.
Introduction

Aphasia is a neurological language disorder that is acquired by injury to the brain, which affects both expressive and receptive language (Hallowell, 2017). Atypical patterns of speech pause are a common impairment for people with aphasia and include any period of time within or between an utterance in which a person is not producing language. Other individuals with various communication disorders exhibit atypical speech pause patterns as well. However, little research has been done to quantify speech pause in the speech of individuals with no known speech or language disorder. Previous research has commonly adopted 250 ms as the standard threshold measurement to differentiate what qualifies as a typical versus nontypical speech pause between utterances; however, empirical evidence for this cutoff is lacking (Oliveira, 2002). In addition, there is limited information regarding how typical speakers use between utterance pauses in their production of speech. By gathering normative data regarding pause in typical speakers, clinicians are provided with a diagnostic marker to use when assessing clients with aphasia and other language disorders.

Although at first glance speech pauses may seem like a lack of expressive communication, research indicates that they can have communicative effects on both the linguistic content being conveyed and the speaker’s perceived qualities, including likability and communicative effectiveness. There are different types of pauses that can accomplish this purpose. Filled pauses include moments when the speaker continues phonating between words. For example, filled pauses might include prolongation of the first sound of an upcoming word or the insertion of a filler word such as “um” or “uh.” Silent pauses are moments when the speaker stops producing speech sounds of any kind when communicating (Price, 2021).
Pauses can also differ in where they happen to be located in an individual’s speech, occurring between or within utterances. Between utterance pauses are typically longer than pauses that occur within utterances and might aid in marking sentence boundaries or the beginning or end of a conversational turn. Alternatively, within utterance pauses occur at times when the speaker pauses between words (e.g., “Hello, how are… (pause) …you?”). These pause types are dynamic and vary from speaker to speaker. The reason for this dynamic variation in pauses might depend on several speaker factors including speech rate, the emotion of the speaker, or the nature of the message being relayed (Price, 2021). The pattern of pauses in speech may vary depending on the speaker’s intended task. For instance, the speaker’s pause behavior may differ when reading a text aloud compared to describing a picture, as demonstrated in research conducted by Zitting (2018).

**Linguistic Factors**

Speech pause is used by typical speakers to accomplish several linguistic functions. Firstly, syntactic boundaries can be marked by using speech pause at the end of a sentence, or even in between clauses. Esposito (2006) and Esposito et al. (2007) recorded narrative samples from children and examined the function of silent pause. They determined that longer speech pauses were utilized to indicate the syntactic boundaries of the children’s utterances. In other words, a speaker might pause longer in between utterance clauses, sentences, or even paragraphs, whereas shorter pauses were used by the young speakers when expressing complex or novel information.

Another linguistic function of speech pause is to mark an increase of information. In a study by Goldman-Eisler (1958), researchers examined speech pause and its correlation with differing levels of novel semantic information among participants. They found that fluent
speakers who exhibited a decrease in speech pause expressed less new information, with lower variability in lexicon. Conversely, less-fluent speakers who had higher levels of speech pause tended to provide more novel information with an increase in vocabulary size.

Speech pauses can serve as markers that separate different narrative elements. Esposito (2006) conducted a study on story retell narrations of 10 female and four male children and found that pauses longer than half a second were used by the children to systematically signal changes in scene, time, and event structures. For instance, a speaker telling a story might pause longer between two scenes to indicate a shift in the narrative. These findings suggest that speakers use pauses strategically to communicate important information to their listeners, and clinicians can use this knowledge to better understand the narrative abilities of their patients. By analyzing a patient’s use of pauses in storytelling, clinicians can gain insight into the patient’s ability to structure and organize their narratives.

**Perceptual Factors**

Speech pause heavily influences a listener’s perception of the speaker’s language in a number of different ways (Bilá & Džambová, 2011). One aspect might be how fluent a speaker is perceived by listeners to be according to their speech pause patterns (Brennan & Williams, 1995). Price (2021) indicated that an increase in speech pause is associated with a decreased perception of speaker likability and communicative effectiveness. This study also found that ratings of listener likability decrease when an extended pause is located prior to the subject or object of an utterance.

Speech pause can also be used by a speaker to express their intended emotion to the listener. In a study by Roberts and Francis (2013), 318 participants were asked to listen to various telephone conversations and rate the recordings based on the perceived enthusiasm of the
speaker. When the speakers paused for longer than 600 ms, participants perceived them as being less enthusiastic. The researchers concluded that speakers can impact a listener’s perception regarding the emotion of their communication by changing how long they pause during their speech.

Bilá and Džambová’s (2011) study focused on the use of pauses in discourse during a conversation about hotels, comparing monolingual and bilingual speakers. L2 speakers were perceived as less fluent but showed more planning, programming, and production time compared to L1 speakers, suggesting that pauses can provide listeners with additional processing time as the speaker plans and programs their speech. Longer pauses, particularly those over 200 ms, were related to syntactic aspects of planning and the lexico-semantic system. Overall, the use of pauses in speech can be seen as an indicator of the speaker’s planning and programming ability. While longer pauses may indicate more careful planning, they may also be seen as less fluent. Shorter pauses, on the other hand, are perceived as more natural and spontaneous. Therefore, the perception of fluency in speech can be influenced by the length of pauses used by the speaker.

Task Types

Tavakoli and Foster (2008) focused on the effect of narrative type on second language (L2) learner output. The researchers found that task type significantly affected the frequency and duration of speech pauses in L2 learners’ narratives. Specifically, L2 learners produced more pauses in response to a picture prompt compared to a written prompt. The authors suggested that task design can influence L2 learners’ fluency and that it is important to consider task type when designing speaking assessments for L2 learners. This finding is relevant to the use of speech pauses, as it suggests that task type can influence the frequency and duration of pauses in speech.
Zitting (2018) investigated perceptual proficiency ratings of obstruent productions in L2 learners of English, as a function of speech task type, word position, and listener expertise. The findings indicated that the task type and word position affected the listeners’ perception of the L2 learners’ speech proficiency. Specifically, listeners rated speech produced in a reading task as more proficient than speech produced in a spontaneous speech task. Furthermore, speech sounds produced at the beginning of a word were rated as more proficient than speech sounds produced at the end of a word. The study’s findings suggest that different speech tasks can influence the perception of L2 learners’ speech proficiency and that speech pauses may be more frequent in spontaneous speech tasks.

**Pause Length**

Previous research has focused on studying and analyzing speech pauses; however, there is currently no universal standard for categorizing pause lengths. As a result, different studies have used different methods for categorizing pauses. Bilá and Džambová (2011) proposed a categorization system for pause lengths that has gained some attention in the field. The Bilá and Džambová’s categorization system includes seven categories, ranging from no pause to extremely long pauses, and can be used to analyze speech samples from individuals with and without language disorders. According to this model, a typical speaker would fall within the “short pause” category, while speakers with aphasia may exhibit longer pauses falling within the “long” or “very long” categories as described below:

- Zero pause or extremely short pause (≤ 50 ms)
- Very short pause (50 ms – ≤ 100 ms)
- Short pause (100 ms – ≤ 300 ms)
- Normal/optimal (300 ms – ≤ 1350 ms)
• Long pause (1,350 ms – ≤ 2,200 ms)
• Very long pause (2,200 ms – ≤ 2,800 ms)
• Extremely long pause (≥ 2,800 ms)

Alternatively, clinicians may use a simpler three-tier model to assess patient speech pauses. Campione and Veronis (2002) suggest classifying pauses as brief (50 ms – ≤ 200 ms), medium (200 ms – ≤ 1,000 ms), or long (≥ 1,000 ms). In general, it would be typical for speakers to have brief or medium pauses, with the occasional long pause. However, patients with nonfluent aphasia may exhibit a higher proportion of long pauses.

**Speech Pause in Individuals With Aphasia**

Considering that speech pause is associated with linguistic and perceptual functions of typical communication, individuals with communication disorders have been found to exhibit atypical patterns of speech pause. This is the case for many people with aphasia, who often display an increase in both the duration and frequency of pausing, especially for speakers with nonfluent types of aphasia. Recently, at Brigham Young University, researchers gathered quantitative data that measured speech pause in people with fluent and nonfluent aphasia. Each pause, both between and within utterances, was examined in 15 people with fluent aphasia, and five people with nonfluent aphasia. The mean pause length was longer for those with nonfluent aphasia than those with fluent aphasia (Thomas, 2021).

DeDe and Salis (2020) also examined speech pause in people with latent and anomic aphasia, compared to individuals with typical communication abilities. Speech recordings were collected from 30 participants who were asked to retell the story of Cinderella. The speech pause patterns from the recorded samples were then analyzed using acoustic software. The researchers found that the nature of the speakers' silent pauses did not differ in people with
anomic versus latent aphasia but did differ in people with either anomic or latent aphasia compared to typical speakers.

It is important to note that the biological mechanism for atypical patterns of speech pause in people with aphasia is still not fully understood. Extended pausing may be the result of damage to the speech and language centers of the brain but may also be due to other concomitant impairment of an individual’s cognitive, respiratory, or muscular systems. Huber et al. (2012) found that when compared to age and gender matched control participants, older adults with neurodegenerative disease produced shorter utterances, had a greater percentage of breaths at minor boundaries, and a smaller percentage of breaths at major boundaries. It was concluded that these changes in respiration may impact an individual’s speech pause patterns during expressive language.

Formal Assessment of Pause in Aphasic Patients

Additional insight into the typical and atypical patterns of speech pause has the potential to impact a clinician’s ability to diagnose a patient with specific types of aphasia. Clinicians should use assessment batteries that not only evaluate a client’s ability to produce and comprehend words, but also consider the significance of prosodic and paralinguistic information in expressing meaning and feeling during communication. This approach is crucial in ensuring that everyday assessments accurately reflect a client’s communication abilities.

One commonly used aphasia assessment is the Quick Aphasia Battery (QAB) developed by Wilson et al. (2018). The QAB utilizes eight different subtests: (a) level of consciousness, (b) connected speech, (c) word comprehension, (d) sentence comprehension, (e) picture naming, (f) repetition, (g) reading aloud, and (h) motor speech. Although the clinician collects these speech samples, the QAB does not have normative data from typical speakers that assist the clinician in
evaluating prosodic speech elements, such as pause, in the speech of people with aphasia. The clinician must solely rely on their subjective opinion regarding the patient’s speech pause frequency and duration.

The Reading Comprehension Battery for Aphasia-Second Edition (LaPointe & Horner, 1998) measures reading comprehension abilities in people with aphasia. This test assesses many aspects of reading such as word order, factual versus inferential reading, and synonym recognition (LaPointe & Horner, 1998). This test also provides a range of activities surrounding reading. However, it does not provide a way for the clinician to rate the person’s speech pause. Although these aphasia batteries take a cursory approach in evaluating pause in people with aphasia, there is relatively little data to make quantitative judgements and examine how that compares to the typical population across various speech tasks.

A survey by Cherney et al. (2018) found that the Western Aphasia Battery (WAB-R; Kertesz, 2006) is one of the more commonly used assessments for the identification of aphasia. The WAB includes a number of subtests that evaluate (a) spontaneous speech, (b) naming and word finding, (c) writing, (d) apraxia, (e) repetition, etc. This test also provides a picture description task that is intended to elicit a spontaneous speech sample from the client to assess different aspects of their speech and language. In the manual of the Western Aphasia Battery, there is a section under the picture description task that provides an aid for the clinician to describe the client’s speech more readily on a subjective level. This manual is a very helpful and comprehensive way for clinicians to rate their client’s speech sample on a scale of zero through 10, however, there is no objective and standardized way for a clinician to rate whether their client is fluent or not based on their speech pause frequency and lengths (Rosen et al., 2010).
Study Purpose

The purpose of this study was to determine how adult speakers with no known speech or language disorder use speech pause and prolongations in their communication, focusing on both filled and unfilled pauses between speakers’ utterances. Specifically, the research questions addressed in this study are as follows:

1. How long do typical speakers pause between their utterances when producing a spontaneous speech sample elicited by completing a picture description task?
2. How frequently do speakers use relatively short pauses between their utterances compared to more extended pauses in a spontaneous speech task?
3. Does the relative length and number of pauses between utterances change as a function of a speakers’ gender or the location within a speech sample?
4. How frequently do typical speakers produce speech sound prolongations between their utterances when describing an everyday picture.

Method

Participants

Audio samples were collected from 60 adults who had no reported history of speech, language, or cognitive impairment. The participants were American English-speaking adults between 20 and 31 years of age ($M = 23.11$ yrs.). Thirty-three participants identified as male and 27 as female. A specific listing of the demographic characteristics of the speakers is included in Table 1. All the participants read and signed a consent form approved by the Institutional Review Board at Brigham Young University prior to having their speech recorded (IRB Study #2022-087).
Speech Task

Participants were recorded while completing a picture description task, drawn from a commonly used assessment battery for the diagnosis of adult language disorders and aphasia. In this task, participants were shown the picnic scene from the Western Aphasia Battery and asked to describe as much as they can for approximately one to two minutes of time. The picture elicitation task included a scene of a group of people, presumably a family, vacationing near the beach (Kertesz, 2006).

Speech Recordings

Participants were recorded using a USB microphone (Blue Yeti) connected to a PC laptop for storage and subsequent editing. Samples were recorded at a rate of 44.1 kHz and 24-bit quantization. The recorded sound files from the participants were filtered to remove any extraneous noise or electronic hum below 60 Hz using Adobe Audition (Adobe Systems, 2003).

C-Unit Segmentation

Speech samples were also segmented into C-units. A C-unit is an independent or main clause along with any modifiers or subordinate clauses. Rules for segmenting these C-units were obtained from Salt Software conventions (Savaldi-Harussi & Soto, 2016). A transcript of each recording was created to assist in dividing the participant recordings into C-units and to aid in the identification of pause locations and prolongations.

Acoustic Measurements

This study focused on filled and unfilled pauses located between utterances or C-unit segments and audible prolongations at the beginning of words. Only between utterance pauses were measured, not within utterance pauses. As shown in Figure 1, the pause locations and durations were marked using a “text grid” which overlays the acoustic signature of the original
sound file using Praat analysis software (Boersma & Weenink, 2022; Öktem et al., 2021). The measured pause periods were computed to the closest millisecond. During a speech recording, any syllables or sounds produced by the participants that were unintelligible or considered to not add to the content of an intended message (e.g., um, ah, like,) were considered a filled pause and were marked on the text grid for subsequent analysis, according to Salt Software conventions (Savaldi-Harussi & Soto, 2016). This also included any laughing, humming, or non-verbal sounds.

Mean pause durations were calculated at different locations within the speech samples to determine if speakers’ pause patterns would change depending on how long the participant had been speaking. Thus, pause values were calculated and averaged for the initial 50%, final 50%, and total duration of the speech samples. In addition, the frequency of pauses of differing lengths were calculated and normalized by the number of C-units. The pause lengths were categorized into pauses of brief (50 ms – ≤ 200 ms), medium (200 ms – ≤ 1,000 ms), or long (≥ 1,000 ms). The mean number of prolongations, irrespective of length, were also averaged across the number of C-units. Frank prolongations were marked using subjective perceptual judgement.

Reliability

To measure the reliability of the filled and unfilled pause values, 10% of all recordings were reanalyzed by a second rater and then correlated with the original rater’s measurements. The recordings to be reanalyzed were randomly chosen across speaker pause type. The first- and second-rater measurements of the initial and ending time values for each pause had a Pearson correlation of $r = .99$, $p < .001$ with a mean absolute difference between the two raters’ values of 5 ms for the initiation portion of each pause and 4 ms when marking the end of the pause segments.
Statistical Analysis

The speech pause dependent measurements collected in this study were descriptively reported in terms of the mean pause duration and standard deviation as a function of the type of pause (silent vs filled), speaker gender, and location within the speech sample. The frequency of occurrence for the silent and filled pauses were categorized by duration and reported as an average of overall duration of the sample and number of C-units. The frequency of prolongation was reported as an average of overall sample duration and number of C-units, irrespective of the length of the prolongation. Inferential analysis of the data involved a repeated-measures analysis of variance (ANOVA) with two within-subject factors of task type and utterance position (between C-unit utterances).

Results

Overall, the speech samples were found to have a variation in speech rate, word count, and number of C-units, but with a relatively consistent duration across the samples. The mean speech rate for the participants was 150.08 words per minute (WPM), with a range from 79.08 to 214.46 WPM. The overall mean duration of the participant samples was 55.86 seconds, with a range from 29.685 to 94.325 seconds. The mean number of C-units was 14.95. The mean word count was 137.63, with a range from 49 to 283 words per participant sample.

Silent Pause Measures

The specific mean and standard deviation values for silent pauses found at the beginning, end, or total duration of the samples are listed in Table 2, organized as a function of the speaker gender. In addition, Table 2 includes the frequency of silent pauses of a particular length normalized by the number of C-units and overall duration of each sample.
Differences between the measures as a function of speaker gender and sample location were inferentially examined using an analysis of variance (ANOVA) statistical test. Findings from the ANOVA revealed significant differences between the silent pauses found in the initial and final positions of the speech samples, \( F(1,58) = 9.96, p = .003, \eta^2_p = .15 \). For male speakers, the initial mean silent pause duration was 80 ms, and 100 ms in the final position. For female speakers, the initial mean silent pause duration was 50 ms, and 80 ms in the final position. Therefore, the average initial silent pause duration was lower than the average final silent pause duration across all participants. No statistical differences were found between male speaker and female speaker samples. In addition, no other main or interaction effects were found to be significant for the silent pause measures.

As shown in Figures 2 and 3, overall frequency of total silent pause ranged from 45.11% to 48.57%. The brief silent pause frequency ranged from 8.34% to 9.92%. The medium silent pause frequency ranged from 30.36% to 32.20%. The long silent pause frequency ranged from 3.00% to 9.87%. There was a higher number of medium silent pauses compared to brief or long pauses. No differences were found between male and female speakers. Error bars indicate the standard deviation.

**Filled Pause Measures**

The specific mean and standard deviation values for filled pauses found at the beginning, end, or total duration of the samples are listed in Table 3 as a function of the speaker gender. In addition, Table 3 also includes the frequency of filled pauses of a particular length normalized by the number of C-units and overall duration of each sample.

Differences between the measures as a function of speaker gender and sample location were inferentially examined using an analysis of variance (ANOVA) statistical test. Findings
from the ANOVA revealed significant differences between the filled pauses found in the initial and final positions of the speech samples, $F(1,58) = 4.70, p = .034, \eta^2_p = .08$. For male speakers, the initial mean filled pause duration was 70 ms, and 120 ms in the final position. For female speakers, the initial mean filled pause duration was 70 ms, and 90 ms in the final position. Therefore, the average initial filled pause duration was lower than the average final filled pause duration across all participants. No statistical differences were found between male speaker and female speaker samples. In addition, no other main or interaction effects were found to be significant for the filled pause measures.

Figures 4 and 5 presents data on filled pause (FP) duration and frequency measures across sample location and speaker gender. In terms of overall frequency, the total filled pause frequency ranged from 19.15% to 19.93%. The brief filled pause frequency was zero for all sample locations and speaker genders. The medium filled pause frequency ranged from 3.40% to 5.17%. The long-filled pause frequency ranged from 14.76% to 15.75%. There was a higher percentage of long filled pauses compared to brief or medium filled pauses. There were no statistical differences between male and female speaker samples for filled pause measures.

**Prolongation Measures**

The frequency of prolongations was calculated, normalized by the number of C-units. In addition, these prolongation frequencies were calculated between male and female speakers for the initial 50% of the samples, final 50% of the samples, and the total duration of the speech sample. The specific mean and standard deviation values of these measures are listed in Table 4.

Differences between the measures as a function of speaker gender and sample location were inferentially examined using an analysis of variance (ANOVA) statistical test. Findings from the ANOVA revealed significant differences between the prolongations found in the initial
and final positions of the speech samples, $F(1, 58) = 0.735, p = .395, \eta^2_p = .013$. As shown in Figure 6, for male speakers the initial mean prolongation frequency was 3.68% and 10.84% in the final position. For female speakers, the initial mean prolongation frequency was 8.05%, and 7.34% in the final position. Therefore, the average initial prolongation frequency was lower than the average final prolongation frequency for males. Conversely, for females, the average initial prolongation frequency was higher than the average final prolongation frequency.

**Discussion**

The general purpose of this study was to examine and quantify speech pause patterns in speakers of American English. By using the picture description task from the WAB-R, it is hoped that findings from this research might provide clinicians with a greater ability to assess the prosody of patients with communication disorders. A discussion of the findings of this study is organized according to the research questions listed below.

**Pause Duration**

The average duration for silent pauses between utterances was approximately 80 ms. Furthermore, the average duration for filled pauses between utterances was approximately 85 ms. These mean pause lengths are lower than the current standard threshold of 250 ms often used to assess typical versus atypical pause durations (Oliveira, 2002). However, it should be noted that the present study averaged the mean pause values by the total number of C-units in the speech sample, thus where a pause was not present zero pause time was average into the statistical analysis for mean duration. Given these findings, it may be appropriate to consider lowering the standardized threshold for what constitutes a typical between-utterance pause. It would be of value to conduct perceptual research to determine if listeners attend to pause durations of 200 ms
or lower. This information would be of value in order to more accurately distinguish between typical and atypical patterns of pause durations.

Bilá and Džambová (2011) found that silent pauses that are shorter than 200 ms can be challenging to distinguish from occlusives or the natural breaks between utterances where a pause is not present. Additionally, Bilá and Džambová (2011) reported that the proportion of very brief pauses (less than 200 ms) ranged from 2.9% to 17.9%. The same study did not report an average pause duration. Direct comparisons between the current study and Bilá’s study are limited by methodological differences. Specifically, Bilá’s study focused on silent pauses only, while the current study examined both silent and filled pause durations. Overall, the current study adds to the growing body of literature on speech pause durations in typical speakers, suggesting that both silent and filled pause durations between successive C-units tend to be relatively short.

The ability to produce short pauses is a hallmark of efficient expressive language. This could be due to several factors, including a more automatic and effortless production of speech, as well as a greater ability to plan and execute language at a faster pace. Short pauses may also facilitate smoother communication, as they allow for more rapid turn-taking and a more natural conversational flow. However, it is important to note that the exact reasons for the prevalence of short pauses in typical speakers is still unclear, and further research is needed to fully understand this phenomenon.

**Pause Frequency**

In a spontaneous speech task for this study, for filled and silent pauses combined, speakers tended to use shorter (brief and medium length) pauses more frequently than extended pauses. The percentage of medium silent pauses was greater than the percentage of brief and
long pauses. The data from the study revealed that brief filled pauses were very limited across all participants. This was due to the inherent difficulty in producing a sound or non-content word (e.g., um, uh) that is less than 200 ms in duration. The percentage of long filled pauses was greater than the percentage of brief or medium pauses.

One possible explanation for the prevalence of longer filled pauses is that these aspects of speech may be a natural aspect of speech production and may serve to facilitate the organization and planning of speech. Longer pauses allow speakers to plan upcoming content, signal the end of a thought or idea, and provide a brief break for listeners to process information, thereby maintaining a more fluid and continuous speech flow. Longer pauses may also be used for more dramatic effect or to signal a significant change in topic or direction, especially when under cognitive load. Khawaja et al. (2007) found that longer pauses, slower speech rate, and higher pitch were all associated with higher levels of cognitive load, indicating that pause duration is a significant speech feature that could be used to measure cognitive load. It is worth noting that their study did not find a minimum duration for a pause to be considered deliberate or significant. These findings could inform the development of new methods for measuring cognitive load in individuals with language impairments, such as aphasia, which could lead to better assessment and treatment of language disorders. The complex interplay between cognitive, linguistic, and social factors that influence how speakers plan, produce, and communicate language likely affects the frequency and duration of pauses in speech.

**Gender and Pause Location**

The data does not suggest any significant differences across speaker gender when it comes to pause patterns. This study did not examine other known prosodic differences between men and women such as pitch and resonance. While speech pause patterns do not differ between
males and females, Price (2021) examined the perceptual differences between male and female pause patterns. She found that female speakers were “rated to have decreased communicative effectiveness and to be less likable than the male speaker across all pause durations” (p. 22). In the assessment process, it might be important for clinicians to objectively examine pause patterns for each patient and not allow subjective data to cloud their judgment. Alternatively, this information could encourage clinicians to focus on decreasing atypical pause patterns in their female patients compared to their male patients knowing that pausing takes a greater toll on their communicative effectiveness.

While there was no discrimination between male and female speakers’ pause patterns, there was a significant difference in the location of the pause. The data shows that both pause frequency and mean pause length increase in the latter half of the speech sample compared to the initial half. This finding suggests that there might be underlying factors that affect a speaker's ability to produce sentences at a fast rate, such as cognitive resources. As speakers continue to talk, they might become mentally fatigued which can lead to longer pauses and slower speech. Additionally, in the context of the picture description task, speakers might run out of things to say, which could also contribute to the increase in pause length and frequency.

From a clinical perspective, these findings are useful for speech-language pathologists who work with patients who have language disorders. By understanding the typical pause patterns of fluent speakers, clinicians can identify potential disparities in their patients' language production. For example, when a clinician is evaluating a person with aphasia, they might compare the first and second half of a speech sample. This information can help clinicians develop targeted interventions to improve their patients' language abilities.
Prolongation Frequency

Based on the data, it appears that both male and female speakers produce speech sound prolongations during a spontaneous speech task describing an everyday picture, with males demonstrating a frequency of 7.26% and females demonstrating a slightly higher frequency of 7.69%. Interestingly, the distribution of these prolongations differed between the genders, with males exhibiting a higher frequency in the final position of the sample (10.84%) compared to the initial position (3.68%), while females demonstrated a higher frequency in the initial position (8.05%) compared to the final position (7.34%). It is possible that these differences in the distribution of prolongations reflect differences in communication styles or social norms between male and female speakers. Overall, these findings suggest that speech sound prolongations are a common feature of spontaneous speech and may have a communicative function, such as indicating emphasis or emotion.

The examination of prolongations is somewhat unique to this study. Typical speakers tend to occasionally produce filled pauses by prolonging a sound usually found in the first speech sound of a word (Clark & Tree, 2002). Typical speakers produce speech sound prolongations at a relatively low frequency when describing a picture. This low frequency of speech sound prolongations in typical speakers may be because most people have learned to speak fluently and without interruptions by the time they reach adulthood. However, speech sound prolongations can be more frequent in individuals with speech disorders or developmental disabilities. Additionally, speakers may produce more sound prolongations in situations where they are nervous or under stress, such as during public speaking or job interviews.
Study Limitation and Future Research

This study provides insights into speech pause durations and frequencies but has several limitations. It only considers pauses between utterances, not within, and involves a sample of 60 speakers without language disorders, focusing solely on pause duration and frequency. Future research should include a larger participant pool and examine pauses within utterances for more comprehensive data.

The study does not take in factors such as age, education, or socioeconomic status, and does not explore variations across different populations or speech tasks. It also exclusively uses Praat acoustic analysis software, leaving room for exploration of alternative methodologies (Boersma & Weenink, 2022; Öktem et al., 2021).

While this research finds speakers favor shorter pauses, and there's a higher frequency of prolongations at the end of speech samples, it does not examine the reasoning behind these patterns. Future studies should investigate these areas further, including potential influences of language complexity and speaker's cognitive load.

The study confines its context to a picture description task, leaving out other speech tasks such as reading, expository speech, and narrative retell. Future research should diversify the tasks and study the impact of cognitive load on speech pause in these varied contexts.

Lastly, despite providing valuable insights, the study is limited in its scope. Future research should address these limitations, explore a broader range of factors, speech tasks, and populations for a more comprehensive understanding of how speech pause durations and frequencies vary.
Conclusion

Despite these limitations, the information provided in this study offers insights and enhanced understanding for clinicians in the assessment of speech pausing. Notably, one of the key characteristics of individuals with aphasia, particularly nonfluent aphasia, is atypical patterns of pause. Thus, the knowledge gained from this study may be relevant for the evaluation and management of individuals with aphasia. Clinicians might utilize the results of this study to compare the typical patterns of pause in healthy individuals with the pause patterns present in speech samples of patients with aphasia. Such a comparison may be useful for identifying atypical pause patterns in patients and distinguishing between fluent and nonfluent aphasia.
References


### Table 1

**Identified Gender and Speech Sample Characteristics for Adult Participants**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Speech Rate</th>
<th>Word Count</th>
<th>Duration</th>
<th>Number of C-units</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>F</td>
<td>165.94</td>
<td>104</td>
<td>37.605</td>
<td>11</td>
</tr>
<tr>
<td>02</td>
<td>F</td>
<td>130.98</td>
<td>131</td>
<td>60.007</td>
<td>14</td>
</tr>
<tr>
<td>03</td>
<td>M</td>
<td>160.33</td>
<td>164</td>
<td>61.375</td>
<td>15</td>
</tr>
<tr>
<td>04</td>
<td>F</td>
<td>137.59</td>
<td>140</td>
<td>61.05</td>
<td>15</td>
</tr>
<tr>
<td>05</td>
<td>M</td>
<td>149.75</td>
<td>158</td>
<td>63.307</td>
<td>14</td>
</tr>
<tr>
<td>06</td>
<td>M</td>
<td>214.46</td>
<td>239</td>
<td>66.865</td>
<td>21</td>
</tr>
<tr>
<td>07</td>
<td>M</td>
<td>186.54</td>
<td>181</td>
<td>58.218</td>
<td>16</td>
</tr>
<tr>
<td>08</td>
<td>M</td>
<td>112.76</td>
<td>128</td>
<td>68.112</td>
<td>13</td>
</tr>
<tr>
<td>09</td>
<td>M</td>
<td>160.66</td>
<td>89</td>
<td>33.238</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>180.02</td>
<td>283</td>
<td>94.325</td>
<td>38</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>151.22</td>
<td>184</td>
<td>73.007</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>149.40</td>
<td>109</td>
<td>43.775</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>147.35</td>
<td>149</td>
<td>60.673</td>
<td>22</td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>174.36</td>
<td>179</td>
<td>61.597</td>
<td>19</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>79.08</td>
<td>79</td>
<td>59.941</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>188.47</td>
<td>202</td>
<td>64.309</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>158.03</td>
<td>233</td>
<td>88.464</td>
<td>26</td>
</tr>
<tr>
<td>18</td>
<td>M</td>
<td>123.00</td>
<td>165</td>
<td>80.49</td>
<td>10</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>118.81</td>
<td>92</td>
<td>46.461</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>172.22</td>
<td>138</td>
<td>48.078</td>
<td>13</td>
</tr>
<tr>
<td>21</td>
<td>M</td>
<td>152.25</td>
<td>213</td>
<td>83.939</td>
<td>16</td>
</tr>
<tr>
<td>22</td>
<td>M</td>
<td>128.03</td>
<td>98</td>
<td>45.629</td>
<td>7</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>158.03</td>
<td>141</td>
<td>53.535</td>
<td>15</td>
</tr>
<tr>
<td>24</td>
<td>M</td>
<td>140.71</td>
<td>84</td>
<td>35.818</td>
<td>7</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>161.14</td>
<td>96</td>
<td>35.745</td>
<td>12</td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>186.90</td>
<td>203</td>
<td>65.168</td>
<td>19</td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>124.27</td>
<td>152</td>
<td>73.39</td>
<td>16</td>
</tr>
<tr>
<td>28</td>
<td>F</td>
<td>177.98</td>
<td>109</td>
<td>36.746</td>
<td>9</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>170.39</td>
<td>94</td>
<td>33.1</td>
<td>8</td>
</tr>
<tr>
<td>30</td>
<td>M</td>
<td>150.39</td>
<td>220</td>
<td>87.813</td>
<td>18</td>
</tr>
<tr>
<td>31</td>
<td>F</td>
<td>142.10</td>
<td>178</td>
<td>75.158</td>
<td>17</td>
</tr>
<tr>
<td>32</td>
<td>F</td>
<td>150.85</td>
<td>117</td>
<td>46.536</td>
<td>11</td>
</tr>
<tr>
<td>33</td>
<td>M</td>
<td>167.91</td>
<td>96</td>
<td>34.305</td>
<td>10</td>
</tr>
<tr>
<td>Subject</td>
<td>Gender</td>
<td>Speech Rate</td>
<td>Word Count</td>
<td>Duration</td>
<td>Number of C-units</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>34</td>
<td>F</td>
<td>129.86</td>
<td>96</td>
<td>44.357</td>
<td>11</td>
</tr>
<tr>
<td>35</td>
<td>F</td>
<td>163.40</td>
<td>161</td>
<td>59.12</td>
<td>16</td>
</tr>
<tr>
<td>36</td>
<td>F</td>
<td>158.70</td>
<td>216</td>
<td>81.661</td>
<td>17</td>
</tr>
<tr>
<td>37</td>
<td>F</td>
<td>150.56</td>
<td>105</td>
<td>41.845</td>
<td>12</td>
</tr>
<tr>
<td>38</td>
<td>F</td>
<td>184.90</td>
<td>107</td>
<td>34.722</td>
<td>10</td>
</tr>
<tr>
<td>39</td>
<td>M</td>
<td>163.72</td>
<td>81</td>
<td>29.685</td>
<td>9</td>
</tr>
<tr>
<td>40</td>
<td>F</td>
<td>152.78</td>
<td>176</td>
<td>69.121</td>
<td>14</td>
</tr>
<tr>
<td>41</td>
<td>M</td>
<td>91.28</td>
<td>49</td>
<td>32.207</td>
<td>3</td>
</tr>
<tr>
<td>42</td>
<td>M</td>
<td>114.12</td>
<td>91</td>
<td>47.845</td>
<td>10</td>
</tr>
<tr>
<td>43</td>
<td>M</td>
<td>146.50</td>
<td>158</td>
<td>64.712</td>
<td>13</td>
</tr>
<tr>
<td>44</td>
<td>F</td>
<td>136.63</td>
<td>142</td>
<td>62.357</td>
<td>15</td>
</tr>
<tr>
<td>45</td>
<td>F</td>
<td>167.63</td>
<td>162</td>
<td>57.984</td>
<td>11</td>
</tr>
<tr>
<td>46</td>
<td>F</td>
<td>131.69</td>
<td>126</td>
<td>57.406</td>
<td>8</td>
</tr>
<tr>
<td>47</td>
<td>F</td>
<td>165.21</td>
<td>150</td>
<td>54.477</td>
<td>17</td>
</tr>
<tr>
<td>48</td>
<td>F</td>
<td>190.68</td>
<td>116</td>
<td>36.5</td>
<td>13</td>
</tr>
<tr>
<td>49</td>
<td>M</td>
<td>159.70</td>
<td>161</td>
<td>60.49</td>
<td>15</td>
</tr>
<tr>
<td>50</td>
<td>F</td>
<td>141.82</td>
<td>133</td>
<td>56.268</td>
<td>17</td>
</tr>
<tr>
<td>51</td>
<td>F</td>
<td>125.27</td>
<td>123</td>
<td>58.912</td>
<td>12</td>
</tr>
<tr>
<td>52</td>
<td>M</td>
<td>149.92</td>
<td>135</td>
<td>54.029</td>
<td>10</td>
</tr>
<tr>
<td>53</td>
<td>F</td>
<td>168.34</td>
<td>165</td>
<td>58.81</td>
<td>13</td>
</tr>
<tr>
<td>54</td>
<td>F</td>
<td>107.14</td>
<td>151</td>
<td>84.562</td>
<td>12</td>
</tr>
<tr>
<td>55</td>
<td>F</td>
<td>122.63</td>
<td>125</td>
<td>61.161</td>
<td>10</td>
</tr>
<tr>
<td>56</td>
<td>F</td>
<td>167.85</td>
<td>111</td>
<td>39.679</td>
<td>10</td>
</tr>
<tr>
<td>57</td>
<td>F</td>
<td>128.06</td>
<td>113</td>
<td>52.944</td>
<td>11</td>
</tr>
<tr>
<td>58</td>
<td>F</td>
<td>153.16</td>
<td>181</td>
<td>70.906</td>
<td>16</td>
</tr>
<tr>
<td>59</td>
<td>F</td>
<td>152.12</td>
<td>157</td>
<td>61.925</td>
<td>16</td>
</tr>
<tr>
<td>60</td>
<td>M</td>
<td>148.49</td>
<td>115</td>
<td>46.467</td>
<td>14</td>
</tr>
</tbody>
</table>
Table 2

Silent Pause (SP) Duration and Frequency Measures Across Sample Location and Speaker

Gender

<table>
<thead>
<tr>
<th>Measures</th>
<th>Sample Location</th>
<th>Male Speakers</th>
<th>Female Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Pause Duration(^a)</td>
<td>Initial</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>Overall Frequency(^b)</td>
<td>Total</td>
<td>48.57%</td>
<td>16.48%</td>
</tr>
<tr>
<td>Brief(^c)</td>
<td>Initial</td>
<td>4.41%</td>
<td>5.34%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>3.94%</td>
<td>6.99%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>8.34%</td>
<td>8.62%</td>
</tr>
<tr>
<td>Medium(^d)</td>
<td>Initial</td>
<td>15.41%</td>
<td>9.56%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>14.95%</td>
<td>7.88%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>30.36%</td>
<td>11.93%</td>
</tr>
<tr>
<td>Long(^e)</td>
<td>Initial</td>
<td>3.53%</td>
<td>5.85%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>6.33%</td>
<td>8.27%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>9.87%</td>
<td>11.62%</td>
</tr>
</tbody>
</table>

Note. \(^a\)Measured in seconds. \(^b\)Frequency per speech sample. \(^c\)50 ms – ≤ 200 ms. \(^d\)200 ms – ≤ 1,000 ms. \(^e\)≥ 1,000 ms.
Table 3

*Filled Pause (FP) Duration and Frequency Measures Across Sample Location and Speaker*

**Gender**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Sample Location</th>
<th>Male Speakers</th>
<th>Female Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td>Pause Duration&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Initial</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td>Overall Frequency&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Total</td>
<td>19.93%</td>
<td>16.48%</td>
</tr>
<tr>
<td>Brief&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Initial</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Medium&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Initial</td>
<td>2.38%</td>
<td>4.23%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>2.79%</td>
<td>5.82%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>5.17%</td>
<td>7.77%</td>
</tr>
<tr>
<td>Long&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Initial</td>
<td>6.64%</td>
<td>7.11%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>8.12%</td>
<td>9.16%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>14.76%</td>
<td>13.36%</td>
</tr>
</tbody>
</table>

*Note. a Measured in seconds. b Frequency per speech sample. c 50 ms – ≤ 200 ms. d 200 ms – ≤ 1,000 ms. e ≥ 1,000 ms.*
Table 4

*Prolongations Across Sample Location and Speaker Gender*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Sample Location</th>
<th>Male Speakers</th>
<th>Female Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td>Frequency of Prolongations</td>
<td>Initial</td>
<td>3.68%</td>
<td>6.70%</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>10.84%</td>
<td>35.13%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>7.26%</td>
<td>17.58%</td>
</tr>
</tbody>
</table>
Figures

Figure 1

Example of Speech Pause Analysis in Praat

Figure 2

Percentage of Occurrence of Silent Pauses Across Pause Length and Speech Sample Location for Female Speakers
Figure 3

*Percentage of Occurrence of Silent Pauses Across Pause Length and Speech Sample Location for Male Speakers*

![Graph showing the percentage of silent pauses across different pause lengths and speech sample locations for male speakers.](image)

Figure 4

*Percentage of Occurrence of Filled Pauses Across Pause Length and Speech Sample Location for Female Speakers*

![Graph showing the percentage of filled pauses across different pause lengths and speech sample locations for female speakers.](image)
Figure 5

Percentage of Occurrence of Filled Pauses Across Pause Length and Speech Sample Location for Male Speakers

Figure 6

Percentage of Occurrence of Prolongations Across Speech Sample Location for Male and Female Speakers
APPENDIX A

Annotated Bibliography


**Objective:** The objective of this preliminary study was to investigate the functions of silent pauses in first language (L1) and second language (L2) speakers of English and German. They hypothesized that the bilingual speakers would have different speech pause than the monolingual. **Methods:** The study used a qualitative approach, analyzing audio recordings of participants' natural conversations to identify different types of silent pauses and their functions. Participants included L1 and L2 speakers of English and German. **Results:** The study identified four types of silent pauses: planning pauses, cognitive pauses, interactional pauses, and lexical pauses. Planning pauses were the most common type in both L1 and L2 speakers of English and German, and were used primarily to plan upcoming speech. Cognitive pauses were more common in L2 speakers than L1 speakers, and were used to process language. Interactional pauses were used to signal turn-taking and regulate conversational flow, and were more common in L1 speakers. Finally, lexical pauses were used to search for words, and were more common in L2 speakers. Researchers found the following on the perceptual aspects of speech pause: L2 speakers were perceived as nonfluent. L2 speakers showed more planning, programming, and production time than L1 speakers. The segmentation in L1 speech is more natural and spontaneous. **Conclusion:** The study suggests that silent pauses serve different functions in L1 and L2 speakers, and that the type and frequency of pauses may
be influenced by language proficiency and cultural background. Length of pause impacts fluency perception. The more a person pauses, the higher chance they have of being perceived as nonfluent in their speech. *Relevance to current work:* By identifying different types of pauses and their functions, the current study may be able to better understand how speech planning and processing, turn-taking, and word retrieval are affected in aphasia. This study analyzed read speech. They created a table wherein they marked where the speakers paused.


*Objective:* The objective of this study was to investigate how listeners use prosody and filled pauses (i.e., "um" and "uh") to infer the metacognitive states of speakers (i.e., their knowledge or uncertainty about a topic). *Method:* The study used an experimental approach, asking participants to listen to recordings of speakers who used different levels of prosody and filled pauses to convey their knowledge or uncertainty about a topic. Participants were then asked to rate the speakers' knowledge and confidence levels.

*Results:* The study found that listeners used both prosody and filled pauses to infer speakers' metacognitive states. Specifically, speakers who used more rising intonation and fewer filled pauses were perceived as more uncertain and less knowledgeable, while speakers who used more falling intonation and fewer filled pauses were perceived as more confident and knowledgeable. *Conclusion:* The study suggests that prosody and filled pauses can serve as cues to listeners about speakers' metacognitive states, and that speakers may use these cues to manage listeners' perceptions of their knowledge and
confidence. *Relevance to current work:* This article provides insight into how filled pauses are used in speech. By understanding the functions of filled pauses in conveying knowledge and uncertainty, researchers may be able to better interpret speech pause lengths in individuals with aphasia who may struggle with language processing and planning. Listeners might perceive the speech of another person differently depending on whether they are using filled pauses or unfilled pauses. Filled pauses might be more effective when it comes to normal speech, but less effective when it comes to answering questions.


*Objective:* The objective of this survey was to identify and describe the aphasia assessment measures currently used in clinical and research settings. *Methods:* The study used a survey approach, collecting data from practicing speech-language pathologists and researchers who specialize in aphasia assessment. Participants were asked to report which aphasia assessment measures they used most frequently and to provide brief descriptions of each measure. *Results:* The study found that the most commonly used aphasia assessment measures were the Western Aphasia Battery, the Boston Diagnostic Aphasia Examination, and the Comprehensive Aphasia Test. Each measure assessed different aspects of language functioning, including fluency, comprehension, naming, repetition, and memory. *Conclusion:* The study suggests that there are multiple aphasia assessment measures available for clinicians and researchers to use, and that each measure has its
own strengths and weaknesses. Across sites and contexts, the Western Aphasia Battery-Revised (WAB-R) was most administered, followed by the Boston Naming Test (BNT), Cognitive Linguistic Quick Test (CLQT) and Communicative Effectiveness Index (CETI). Relevance to current work: This article provides context for the use of different aphasia assessment measures. By understanding the different aspects of language functioning that are assessed by these measures, researchers may be able to better interpret speech pause lengths in individuals with aphasia and develop more targeted treatment plans. To gather data: our study used examples from the WAB so clinicians can observe differences in people with and without aphasia.


**Objective:** To investigate the effects of aphasia on the temporal and episodic structure of discourse. To measure different speech characteristics, such as silent pause, in people without aphasia, people with latent aphasia, and people with anomic aphasia. **Methods:** Patients with latent aphasia were asked to retell the story of Cinderella. The retellings were analyzed for pause duration and distribution, as well as for the structure of the story in terms of its temporal and episodic components. **Results:** Patients with aphasia displayed significant differences in pause duration and distribution compared to healthy controls and showed a reduced ability to organize the story into its temporal and episodic components. **Conclusion:** Silent pause did not differ in people with anomic versus latent aphasia. But it did differ in people with either anomic or latent aphasia compared to typical speakers. **Relevance to current work:** Pauses that are longer than 200 ms are
related to syntactic aspects of planning and the lexico-semantic system. This study on
temporal and episodic structure of discourse in latent aphasia patients can provide
insights into the effects of aphasia on speech pause lengths and distribution, as well as the
overall organization of speech.

NOLISP 2005 Lecture notes in computer science (pp. 108-115). Springer.
https://doi.org/10.1007/11613107_9

Objective: To investigate the role of pausing in the development of discourse
organization in children. Methods: Speech of children aged 5-7 years as they retold
stories was analyzed to see how they used pausing to indicate changes in topic, shifts in
perspective, and other aspects of discourse structure. Results: The children were able to
use pausing effectively to organize their discourse, and this ability was related to their
overall language proficiency. Conclusion: Pausing is an important cue to discourse
structure, and its use can provide valuable insights into the development of language
proficiency in children.

Relevance to current work: This study on pausing as a cue to discourse structure
in children can provide insights into how speakers use pauses to indicate changes in topic
and shifts in perspective and can be used to inform a study on speech pause lengths in
typical speech. Speech pause is used to separate narrative elements.

**Objective:** To investigate the cognitive and algorithmic issues involved in the use of empty speech pauses. This study sought to examine how children use different pause strategies including filled and unfilled pauses, and phoneme lengthening. **Methods:** The study used a computational approach to analyze empty speech pauses, using a probabilistic model to simulate the cognitive processes involved in the production and perception of speech. Speech from participants was recorded, and pauses were measured and divided into short, medium, and long pauses. **Results:** The study found that empty speech pauses play an important role in facilitating communication, by providing speakers with time to plan and organize their speech, and by allowing listeners to process the information more efficiently. **Conclusion:** Empty speech pauses are an integral part of speech production and perception, and understanding their cognitive and algorithmic underpinnings can provide valuable insights into the mechanisms underlying language use. The longer pause lengths were associated with a change of scene to signal discourse boundaries. **Relevance to current work:** This study on the cognitive and algorithmic issues involved in the use of empty speech pauses can provide insights into the function of pauses in speech.

https://doi.org/10.1080/17470215808416261

Objective: To investigate the effect of predictability on speech production. Methods: Participants were asked to produce a sentence with a target word that was either highly predictable or less predictable given the preceding context. The participants' speech production was analyzed for the presence and duration of pauses. Results: The study found that highly predictable words were produced more quickly and with fewer pauses than less predictable words. Additionally, the study found that the predictability of the target word influenced the duration of pauses that occurred after the word was produced. Conclusion: The predictability of a word in context can significantly impact speech production, including the length and distribution of pauses. Speech pause does show an increase in information in speech. People who use specific words habitually were determined to be more fluent than others who didn’t have these habitual words.

Relevance to current work: This study on the effect of predictability on speech production can provide insights into how contextual factors influence speech pauses. People use speech pause to cognitively plan and prepare to speak.


This text defines aphasia with four elements. It is an acquired, neurologically caused language disorder that affects expressive and receptive language across modalities. It also does not involve psychiatric, motor, or sensory disorders. Fluency in aphasia refers to the
ease and flow of speech production, which can be disrupted in individuals with aphasia.

Fluent speech is characterized by normal prosody, rhythm, and intonation, with words and sentences produced smoothly and without undue effort. In contrast, nonfluent speech is characterized by halting, effortful, and disjointed speech with frequent pauses, hesitations, and word-finding difficulties. Nonfluent aphasia is often associated with damage to the left hemisphere of the brain, particularly in the regions involved in language production, while fluent aphasia is associated with damage to the right hemisphere or the posterior language areas of the left hemisphere. Understanding the nature of fluent and nonfluent speech in aphasia can help clinicians develop appropriate interventions and treatment strategies for individuals with this condition.


Objective: The present study examines the impact of typical aging and Parkinson’s disease (PD) on the relationship among breath pausing, syntax, and punctuation. 

Methods: The study involved analyzing transcriptions of spoken language samples from three groups of participants: healthy young adults, healthy older adults, and individuals with PD. The analysis focused on breath pauses, syntactic complexity, and punctuation. 

Results: The study found that both typical aging and PD were associated with increased breath pausing and decreased syntactic complexity, while punctuation was not significantly affected by either aging or PD. Additionally, the relationship between breath pausing and syntax was found to be weaker in individuals with PD compared to healthy
older adults. *Conclusion:* The study concluded that breath pausing and syntactic complexity are related in spoken language, and that both are affected by typical aging and PD. The weaker relationship between breath pausing and syntax in PD suggests that this could be a potential marker for the disease. There was a greater percentage of breaths, for typical adults, at minor boundaries for typically aging adults compared to people with PD. This did not negatively impact speech intelligibility or naturalness. It is unknown what percentage of speech pause will start to impact intelligibility. *Relevance to current work:* This article suggests that breath pausing is related to other aspects of spoken language, such as syntactic complexity, and that changes in breath pausing could be indicative of neurological disorders like PD. This article gives its own definition of a breath pause. A breath group was defined as all of the words produced on one breath, and utterance length was defined as the number of syllables produced on each breath group. The location of a breath was determined using the sum signal from the Respitrace, which was computed by summing the calibrated RC and AB signals. A breath was defined as a sharp upward deflection in the sum signal.


*Objective:* This study measured pause length, pause frequency, and latency to response under differing levels of cognitive load. Perhaps there is a correlation between pause length/latency and how much cognitive load there is. *Methods:* The study involved analyzing speech recordings of participants performing a task that varied in cognitive
load levels. The analysis focused on various speech features, such as pause duration, speech rate, and pitch. They had a control group, and a dual-task group. For both groups, they were asked to read a passage, and answer questions about it. For the dual-task group, they also wore a headset, and listened to a series of random two-digit numbers. The subjects counted how many numbers they heard. This study also measured their GSR.

**Results:** The study found that pause duration, speech rate, and pitch were all significantly correlated with cognitive load levels. Specifically, longer pauses, slower speech rate, and higher pitch were all associated with higher levels of cognitive load. **Conclusion:** The study concluded that pause duration, speech rate, and pitch are all potential speech features that could be used to measure cognitive load in spoken language. Pause length, pause frequency, and response latency times all increased in the groups that participated in the dual-task (high load). **Relevance to current work:** This study suggests that pause duration is one of several speech features that could be used to measure cognitive load. Additionally, the study's findings could inform the development of new methods for measuring cognitive load in individuals with language impairments, such as aphasia. This article stated how there is not a minimum duration of what constitutes a deliberate or significant pause. If the pause was shorter than 0.3s, it was assumed that it was part of natural speech.

**Objective:** To investigate the role of pause occurrence and pause duration in the signaling of narrative structure. They measured speech duration and speech pause occurrence. Pause is important for oral narratives. **Methods:** The study analyzed a set of spoken narratives and measured the occurrence and duration of pauses in different locations. Researchers differentiated between silent pauses and filled pauses. **Results:** The study found that pause occurrence and duration can be used to signal narrative structure, specifically indicating major and minor boundaries. **Conclusion:** The study suggests that speakers use pauses intentionally to signal narrative structure and that these pauses can be informative to listeners. Longer pauses tend to be more present near the end of a narrative discourse. Pause is important in conversation because not only does it give the speaker some time to formulate their next sentence, but it also provides time for the listener to cognitively process what the speaker is saying. **Relevance to current work:** This study provides evidence for the intentional use of pauses in speech and suggests that the occurrence and duration of pauses can convey information to listeners. This study is relevant in its discussion of measures of speech pause lengths and their communicative function and potential use in speech processing algorithms. This article provides the number 250 ms as the most widely accepted amount of time that a pause takes.


**Objective:** To create a corpus of spoken language with detailed prosodic information for use in speech processing research. **Methods:** The study compiled a large corpus of spoken language with detailed annotations of prosodic features, such as pitch, duration, and
intensity. They implemented three steps. In the first, they segmented speech at the sentence level. They then extracted timing information. They used the Proscript library to annotate the prosodic features. Lastly, they aligned the segments in each of the languages with the corresponding segment. Results: The study produced a large and valuable resource for speech processing research that includes detailed prosodic information. Conclusion: The study demonstrates the importance of using high-quality and detailed speech corpora for speech processing research, particularly in the area of prosody. One of the frameworks was used to make “structured parallel speech data from dubbed movies.” This article introduced a way in which to gather, handle, store, and visualize speech pause data. Using a prosograph is a helpful way to view audiovisual prosody. Relevance to current work: These two programs might be beneficial in recording and measuring speech pause.


Objective: To answer how does pause length within a sentence impact speaker effectiveness, and also the way they are perceived by others. Does the location of the speech pause (before a noun, verb, object, etc…) impact the perception? Methods: The study presented listeners with recordings of speakers using different lengths and locations of pauses and asked them to rate the speakers' likability and communicative effectiveness. 40 participants were used. Results: The study found that shorter pauses were associated with higher ratings of speaker likability and communicative effectiveness, and that longer pauses at the end of sentences were rated as more effective.
than longer pauses in other locations. Conclusion: The study suggests that the length and location of pauses can have an impact on listener perceptions of speaker likability and communicative effectiveness, and that shorter pauses may be preferred in some contexts. As pause duration increased, the communicative effectiveness decreased. Relevance to current work: This study provides insights into the potential communicative function of different pause lengths and locations.


Objective: The objective of this study is to determine the maximum duration of a pause that can occur after a request is made before the listener perceives it as awkward or unnatural. Method: Three hundred and eighteen participants were asked to listen to different telephone conversations. They were then asked to rate the speaker's enthusiasm and willingness on a six-point Likert scale. Results: The results of the study suggest that listeners can tolerate pauses up to 2 seconds after a request before it becomes unnatural. Conclusions: The conclusion of the study is that a temporal threshold of tolerance for silent gaps after requests exists and should be taken into consideration in speech communication. Long pauses of at least 600 ms were perceived as less willing and less enthusiastic. Relevance to current work: There are negative effects with silent elongated speech pause over 600 ms. This study can be relevant to a study that measured speech pause lengths because it provides insights into how pauses can affect the naturalness of speech and how listeners perceive these pauses. By understanding the maximum duration
of a pause that listeners can tolerate, researchers can determine if speech pauses fall within the range of natural pauses or if they are perceived as unnatural by the listener.


**Objective:** The objective of this study is to develop a computer-based method for measuring speech pauses in both normal and dysarthric speech. 

**Methods:** The method used in this study involves analyzing audio recordings of speech and using algorithms to automatically detect and measure pauses. They recorded three minutes of speech from 13 people who have Friedreich’s Ataxia, and 18 typical speakers. 

**Results:** The results of the study suggest that the computer-based method is accurate in detecting and measuring pauses in both normal and dysarthric speech. 

**Conclusion:** The conclusion of the study is that the automatic method of pause measurement can be a useful tool for researchers studying speech pauses in both normal and dysarthric speech. 

**Relevance to current work:** Rosen’s study can be relevant to the current study that measured speech pause because it provides a method for automatically measuring pauses in speech. This can be useful for researchers who want to measure a large number of speech pauses accurately and efficiently. Additionally, the study shows that the automatic method can be applied to both normal and dysarthric speech, which can be useful for researchers studying speech disorders such as dysarthria. Subjective judgement of speech pause is not an infallible measure of fluency.

**Objective:** The objective of this study is to investigate the effect of narrative type on second language performance and output in terms of fluency, accuracy, and complexity.

**Methods:** The method used in this study involves assigning participants to two different narrative tasks, personal recount and story retelling, and measuring their language output in terms of fluency, accuracy, and complexity. **Results:** The results of the study suggest that the type of narrative task can have an effect on second language performance and output, with personal recount tasks leading to more fluent and complex output than story retelling tasks. **Conclusion:** The conclusion of the study is that task design plays an important role in second language learning and teaching, and that different types of tasks can lead to different outcomes in terms of language performance and output. **Relevance to current work:** Tavakoli’s study can be relevant to our study that measured speech pause because it provides insights into how task design can affect language performance and output. By understanding the factors that influence language output, researchers can better understand how pauses are used in second language speech and how they can be measured. Task type has a linguistic impact.


https://scholarsarchive.byu.edu/etd/8939

**Objective:** The objective of this study is to compare the duration of speech pauses in speakers with nonfluent and fluent aphasia and to determine whether speech pauses can
be used as a diagnostic tool for different types of aphasia. 

Methods: 20 participants: 15 with fluent aphasia, and 5 with nonfluent aphasia. Each pause between and within utterances was measured to the millisecond. The method used in this study involves analyzing audio recordings of speech from speakers with nonfluent and fluent aphasia and measuring the duration of speech pauses. 

Results: The results of the study suggest that speakers with nonfluent aphasia have longer and more frequent speech pauses than speakers with fluent aphasia, and that speech pauses can be used as a diagnostic tool for different types of aphasia. 

Conclusion: The conclusion of the study is that speech pauses can provide important information about the type and severity of aphasia, and that they can be a useful tool for diagnosing and treating the condition. 

Relevance to current work: Thomas’ study can be relevant to our study that measured speech pause because it provides insights into how speech pauses can be used as a diagnostic tool for different types of aphasia. By understanding how speech pauses are affected by different types of aphasia, researchers can better understand the underlying mechanisms of the condition and develop more effective treatments. This study also measured speech pause; however, they measured speech pause in people with aphasia. The statistical analysis looks similar. 


https://scholarsarchive.byu.edu/etd/7381

Objective: The objective of this study is to investigate the effect of speech task type, word position, and listener expertise on the perceptual proficiency ratings of obstruent productions in second language learners of English. 

Methods: The method used in this
study involves assigning participants to different speech tasks and measuring the perceptual proficiency ratings of their obstruent productions by expert and non-expert listeners. Speakers were asked to participate in a paragraph reading, retelling a fairytale, and answering open ended questions. Listeners then rated them on a scale from 0 to 100 in clarity. Results: The results of the study suggest that speech task type, word position, and listener expertise can have an effect on the perceptual proficiency ratings of obstruent productions in second language learners of English. Overall, in most phonemes, listeners rated greater clarity in the paragraph reading samples compared to the spontaneous speech samples. Theoretically, the L2 speakers were able to focus on their production rather than the language itself, lending more clarity to their speech. Conclusion: The conclusion of the study is that speech task type, word position, and listener expertise should be taken into consideration when assessing the perceptual proficiency of second language learners of English. Relevance to current work: Zitting’s study can be relevant to our study that measures speech pause because it provides insights into how different factors can affect speech production and perception in second language learners of English. By understanding the factors that influence speech production and perception, researchers can better understand how pauses are used in second language speech.
Appendix B

Consent Form

Consent to be a Research Subject

Title of the Research Study: Communicative Impact of Speech Characteristics - Production Principal Investigator: Shawn Nissen, Ph.D.  
IRB ID#: 2022-087

Introduction
This research study is being conducted by Shawn Nissen, Ph.D., at Brigham Young University to examine the acoustic characteristics of speech while reading, describing a picture, and engaged in everyday conversation. You were invited to participate because you are an adult native speaker of American English.

Procedures
If you agree to participate in this research study, the following will occur:

- You will be asked to complete a short questionnaire that asks about your age in months/years, native language, and if you speak another language.
- You will be asked to read a short paragraph, describe a picture, retell a culturally appropriate fairy tale, and respond to some simple everyday conversational questions. During these tasks your speech will be recorded.
- It is anticipated that your participation will involve 10 minutes of your time.

Risks/Discomforts
There are minimal risks for participation in this study. You may encounter some anxiety having your speech recorded. There is also a small risk that your participation in the study may be known to others by your signing the consent form. In addition, information about your age and language status may be known to others. The consent forms will be kept in a locked cabinet within a locked room to decrease this risk.

In Case of Research Related Injury
BYU makes no commitment to provide financial compensation or free medical care should you be injured as a result of your participation in this research. Nonetheless, in the event of such an injury, after seeking appropriate medical attention, please contact Shawn Nissen at (801) 422-5056 or shawn_nissen@byu.edu.

Benefits
There are no direct benefits to you. It is hoped this study will provide understanding in how to help individuals learn to communicate more effectively.
Confidentiality

All data, including the recordings of your speech, will be kept on password-protected computers in a locked laboratory and only those directly involved with the research will have access to them. The consent forms with the participant signatures is what will be stored in a locked cabinet.

De-identified data from this study may be shared with the research community, with journals in which study results are published, and with databases and data repositories used for research. We will remove or code any personal information that could directly identify you before the study data are shared. Despite these measures, we cannot guarantee anonymity of your personal data.

Data Sharing

We will keep the information we collect about you during this research study for analysis and for potential use in future research projects.

De-identified data from this study may be shared with the research community, with journals in which study results are published, and with databases and data repositories used for research. We will remove or code any personal information that could directly identify you before the study data are shared. Despite these measures, we cannot guarantee anonymity of your personal data.

The results of this study could be shared in articles and presentations, but will not include any information that identifies you unless you give permission for use of information that identifies you in articles and presentations.

Compensation

You will not receive compensation for your participation in this study.

Participation

Participation in this research study is voluntary. You have the right to withdraw at any time or refuse to participate without penalty.

Questions about the Research

If you have questions regarding this study, you can contact the principal investigator Shawn Nissen at (801) 422-5056 or shawn_nissen@byu.edu.

Questions about Your Rights as Research Participants

If you have questions regarding your rights as a research participant contact Human Research Protections Program by phone at (801) 422-1461; or by email: BYU.HRPP@byu.edu

Statement of Consent

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Name (printed): ____________________________ Date:

Signature: ________________________________