



# Effect of hesitation sound phonetic quality on perception of language fluency and accentedness

Tillena Trebon\*, Linguistics

## ABSTRACT

This study investigates the perceptual consequences of nonnative versus native hesitation sounds in evaluating male speech. When the phonetic quality of a hesitation sound is consistent with native speaker hesitation sounds, the hesitation sound is “native.” A hesitation sound with phonetic quality inconsistent with native speaker hesitation sounds is “nonnative.” In Experiments 1A and 1B, participants rated sentences for fluency and accentedness on a Likert scale. In Experiments 2A and 2B, listeners performed a forced choice task to evaluate speech for accentedness and fluency. In Experiments 1A and 1B, hesitation sound phonetic quality did not impact listeners ratings. However, in Experiments 2A and 2B, participants deemed sentences with nonnative hesitation sounds less fluent and more accented compared to those with native hesitation sounds. Results show that the hesitation sound phonetic quality can have perceptual consequences and that the type of task listeners performed to evaluate speech affected accentedness and fluency judgments. This study has important implications for how learners treat pausing when practicing their second language.

---

## 1. INTRODUCTION

Despite popular culture narratives that pausing while speaking should be avoided (see e.g., Cohen, 2012; Riegel, 2018), research shows that pausing is an important part of linguistic communication. Broadly, there are two types of pauses: filled pauses, which are when a speaker makes a sound while pausing, and unfilled pauses, which are silent.

While definitions of native and nonnative speakers vary widely (see e.g., Davies, 2003; Davies & Elder, 2008), native speakers of a specific language are generally considered to be people who learned the language during childhood, and nonnative speakers are people who learned the language after childhood. Both native and nonnative speech contain pauses (Erbaugh, 1987; Goldman-Eisler, 1968). However, nonnative speech has different pausing patterns compared to native speech. Nonnative speech contains more frequent and longer pauses than native speech (Pickering, 1999; Riazantseva, 2001). Nonnative speakers produce more within clause pauses than native speakers (Riazantseva, 2001). Additionally, Erbaugh (1987) found that nonnative speakers display more individual variation in pausing patterns than native speakers. Variable proficiency may influence nonnative pausing patterns; less proficient speakers have longer and

\*Tillena Trebon, an undergraduate double majoring in Linguistics and Spanish, has been a member of the Speech Perception and Production Lab since 2018. For her honors thesis research, reflected in this paper, she has received the VPRI Undergraduate Fellowship, the Robert W. Young Award from the Acoustical Society of America, CURE Conference Travel Award, and the 2021 Undergraduate Linguistics Association of Britain Conference Best Poster Award. Please direct correspondence to [ttrebon@uoregon.edu](mailto:ttrebon@uoregon.edu).

more frequent pauses than higher proficiency speakers (Anderson-Hsieh & Venkatagiri, 1994; Iwashita et al., 2008; Kormos & Dénes, 2004; Riazantseva, 2001; Towell et al., 1996; Trofimovich & Baker, 2006).

Filled pauses, which are the focus of this paper, accomplish crucial linguistic work such as facilitating conversational turn-taking (Clark & Fox Tree, 2002; Maclay & Osgood, 1959), signaling to listeners what a speaker might say next (Brennan & Williams, 1995; Fox Tree, 2001; Watanabe et al., 2008), and giving speakers time to plan utterances (Clark & Fox Tree, 2002; Maclay & Osgood, 1959). Different languages use different sounds for filled pauses; this is described as phonetic quality. English speakers primarily use “uh,” the central vowel [ə] to hesitate (Maclay & Osgood, 1959), with brackets denoting the International Phonetic Alphabet (IPA)’s standardized system used by linguists to identify sounds. In contrast, Spanish speakers primarily use “eh” or [e] to hesitate (Campillos-Llanos & Plá, 2009; Cenoz, 1998; Erker & Bruso, 2017; Roggia, 2012). When the phonetic quality of a hesitation sound is consistent with a hesitation sound used by native speakers, the hesitation sound is “native.” A hesitation sound with phonetic quality inconsistent with a native speaker hesitation sound is “nonnative.”

Pauses also have important perceptual consequences. The perceptual consequences of pause length, frequency, and location in nonnative speech have been well investigated. Longer pauses, more frequent pauses, and unfavorable pause placement (such as within a clause) increase accentedness ratings and decrease fluency ratings of nonnative speech (Bosker et al., 2013; Cucchiariini et al., 2002; Derwing et al., 2004; Kahng, 2018; Kang, 2010; Trofimovich & Baker, 2006; Wennerstrom, 2000).

However, research regarding the phonetic quality of second language hesitation sounds is limited. Studies suggest that second language speakers generally do not produce hesitation sounds with the same phonetic quality as native speakers, although proficiency and speech community are influential factors (Erker & Bruso, 2017; Hlavac, 2011; Rose, 2017). Rose (2017) found that native Japanese speakers learning English did not accurately produce the native English hesitation sound [ə]. However, higher-proficiency learners’ hesitation sounds more closely resembled the acoustic properties of native English hesitation sounds, indicating that mastery of the second language sound system may be important for producing native hesitation sounds. In a study of native Spanish speakers living in the US and learning English, Erker & Bruso (2017) found that native Spanish speakers were more likely to use English hesitation sounds instead of Spanish hesitation sounds when they: (1) frequently interacted with native English speakers, (2) had lived in the US for a longer period of time, (3) and spoke English often.

This paper asks whether the native or nonnative phonetic quality of a hesitation sound impacts listener judgments about fluency and accentedness. This study uses distinct experimental designs to investigate the perceptual consequences of Spanish versus English hesitation sounds in nonnative English speech produced by male speakers.

## 2. OVERALL METHODS

Four speech perception experiments were conducted to investigate how the phonetic realization of a hesitation sound (whether the nonnative speaker uses a native or nonnative

hesitation sound to pause) affects listener judgments about accentedness and fluency. While the term “fluency” has been defined in various ways (e.g. Bosker et al., 2013; Chambers, 1997; Lennon, 1990), in this study, fluency is used as a proxy for language proficiency. As described below, fluency was defined to participants as “how well someone speaks a language.” Two of the four experiments in this study, Experiments 1A and 1B, utilized a Likert scale rating task where participants listened to sentences and rated how accented or fluent the sentences sounded on a scale of 1 to 9. The other two experiments, Experiments 2A and 2B, utilized a forced choice task, presenting participants with pairs of sentences that were identical except that they contained different types of pauses. Participants heard three sentence pairs that had a native and a nonnative hesitation sound, a nonnative hesitation sound and a silent pause, and a native hesitation sound and a silent pause. Participants then chose which sentence of each pair sounded more fluent or more accented. All experiments are described in detail below.

### 3. STIMULI

Because this research was conducted during the COVID-19 pandemic, stimuli were drawn from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSTAR; Bradlow, n.d.). L1 designates a first or native language; L2 designates a second or nonnative language. ALLSTAR is a corpus of L1 and L2 speech that contains over 120 talkers in over 20 languages. Spontaneous speech in the ALLSTAR corpus was elicited via open ended prompts about the talker’s life and personal experiences as well as by asking participants to describe what happened in a series of cartoon pictures. The male voices utilized in the study had spontaneously produced native and nonnative hesitation sounds in the ALLSTAR corpus.

In this study, participants listened to spontaneous speech in English from eight male speakers who spoke Spanish natively and English as a second language. Male speakers and not female speakers were chosen for several reasons. First, studies show that male and female voices can be perceived differently (Boyle, 2015; Klofstad et al., 2012). To avoid factors that could confound the results of this study, and because there was more data in the ALLSTAR corpus available from L1 Spanish and L2 English male speakers than female speakers, only male voices were used to create stimuli. To conceal the research question, stimuli sentences contained both filled and unfilled pauses, and 12 filler sentences were created by identifying relatively short English sentences without pauses. Filler sentences were used only in the Experiments 1A and 1B due to their differing experimental designs compared to Experiments 2A and 2B. The possible effects of using filler sentences in Experiments 1A and 1B and not using filler sentences in Experiments 2A and 2B are addressed in the discussion section of this paper.

Stimuli sentences were created by identifying sentences in spontaneous English speech that contained one filled pause. Forty such sentences were identified. Then, prototypical English and Spanish hesitation sounds for each speaker were identified in English and Spanish spontaneous speech. Using Praat (Boersma & Weenink, 2020), a software program for analyzing and editing speech, either a Spanish hesitation sound, English hesitation sound, or a silent pause were edited into each of the 40 English utterances in the place of the naturally occurring hesitation sound. Thus, each sentence with a naturally occurring hesitation sound yielded three edited stimuli sentences: one with a native English hesitation sound (uh), one with a Spanish hesitation sound (eh) that is nonnative to English, and one with a silent pause. Speaker voices were never mixed;

that is, the hesitation sounds edited into the stimuli sentences always came from the same speaker who produced the sentence. Pause location was never altered from the original pause location produced in spontaneous speech.

All Spanish and English hesitation sounds used in this study had acoustic properties, specifically, F1 and F2 values, consistent with the prototypical English and Spanish vowels used to hesitate. The average F1 and F2 of Spanish hesitation sounds in stimuli were 474.7 Hz (SD = 58.7 Hz) and 1979.9 Hz (SD = 161.7 Hz), respectively, which are similar to those reported for the Spanish hesitation sounds' vowel [e] in Bradlow (1995). The average F1 and F2 of English hesitation sounds in stimuli were 631.5 Hz (SD = 93.9 Hz) and 1343.7 Hz (SD = 129.4 Hz), respectively, which are similar to those reported in Bradlow (1995) for the English [ʌ], a vowel that is perceptually similar to the vowel [ə] used to hesitate in English. Using Praat, stimuli and filler sentences were leveled for intensity (loudness) at 60 dB SPL.

## 4. OVERVIEW OF PARTICIPANTS

For all experiments in this study, no participant reported any speech or hearing impairments. Participants were required to use headphones to listen to speech and complete the experiment in a quiet environment. Participants were either MTurk workers, who were paid up to \$4.10 for their participation, or University of Oregon students from the Psychology and Linguistics Human Subjects Pool, who received credit to satisfy course requirements in return for their participation. Mturk participants were paid using funds provided by the Acoustical Society of America via the Robert W. Young Award for Undergraduate Student Research in Acoustics. Mturk participants and participants from the Psychology and Linguistics Human Subjects Pool were compensated regardless of the content of their responses. Although participant knowledge of linguistics and pausing may have varied, participants were instructed to make judgments based on perception and not based on knowledge. Each study participant only participated in one experiment.

After completing the experiment, participants completed a language background survey that asked about their native language and experience with other languages and nonnative accents. Data gathered in this survey was only used to ensure that participants were native speakers of English.

## 5. EXPERIMENT 1A

### 5.1. EXPERIMENT 1A: PROCEDURE

This experiment was administered online using Qualtrics (Qualtrics, Provo, UT). Over the course of approximately 30 minutes, participants listened to 40 stimuli sentences and 12 filler sentences for a total of 52 sentences in a randomized order and rated each one for fluency on a Likert scale of 1 to 9. Participants were given the following instructions regarding how to rate sentences for fluency:

In this experiment, you will listen to speech and make judgments about how fluent the speaker sounds in English in each sentence. You will rate English fluency on a scale of 1-9. Fluency means how well someone speaks a language.

It is important that you rate each sentence individually even if you think that you've heard the speaker's voice in a previous sentence.

1 = speaker does not at all sound fluent in English in this sentence (beginner in English)

9 = speaker sounds very fluent in English in this sentence (speaks English very well)

To ensure that a participant never heard the same utterance more than once, participants were randomly assigned to one of three conditions. In each condition, a participant heard only one of the three versions of the sentences that contained pauses and 12 filler sentences without pauses. Each condition contained an equal proportion of sentences with English filled pauses, Spanish filled pauses, and silent pauses.

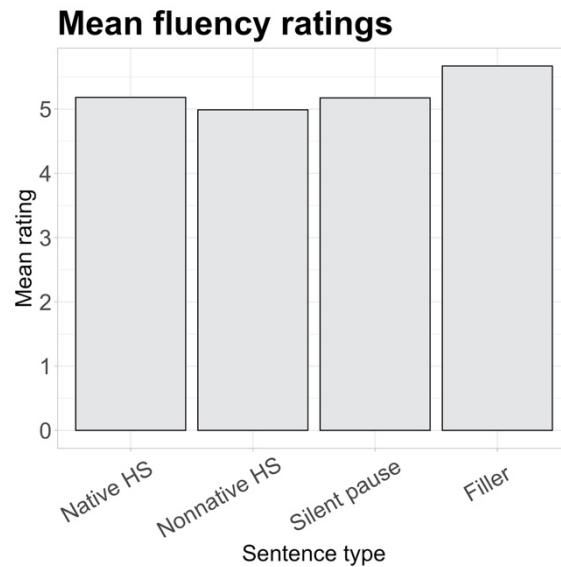
## 5.2. EXPERIMENT 1A: PARTICIPANTS

Fifty native English listeners participated in this experiment. Twenty-nine participants were female, 20 were male, and one participant chose not to disclose their gender. Participants had a mean age of 29.3 years and the age range for participants was 18 to 67 years. Twenty-one participants were Amazon Mechanical Turk (Mturk) workers, and all other participants were University of Oregon students from the Psychology and Linguistics Human Subjects Pool.

## 5.3. EXPERIMENT 1A: RESULTS

Data was analyzed in R (R Core Team, 2020) using linear mixed models analysis. The phonetic realization of the hesitation sound (whether the speaker used the nonnative Spanish "eh" or the native English "uh" to hesitate) did not impact listener judgments about fluency and was not a significant predictor of model fit ( $\chi^2 = 0.0815$ ,  $p = 0.7753$ ). Data analysis also shows that whether the pause listeners heard was filled or unfilled did not influence judgments about fluency ( $\chi^2 = 0.102$ ,  $p = 0.7495$ ). In addition, whether the sentence that listeners heard contained a pause was not a significant predictor of fluency ratings ( $\chi^2 = 1.6311$ ,  $p = 0.2015$ ).

In other words, Experiment 1A results reveal that whether the hesitation sound listeners heard in sentences was native or nonnative did not affect fluency ratings. Additionally, whether the pause listeners heard was filled or silent did not affect fluency judgments, and whether a sentence contained a pause or not did not impact listener ratings. Figure 1 below shows the average fluency ratings for sentences with each pause type. For Figure 1 and all other figures in this paper, HS stands for hesitation sound.



**Figure 1:** Mean fluency ratings for Experiment 1A

## 6. EXPERIMENT 1B

Research shows that nonnative speakers are more likely to produce longer pauses, pause more frequently, and pause within a clause compared to native speakers (Anderson-Hsieh & Venkatagiri, 1994; Iwashita et al., 2008; Kormos & Dénes, 2004; Pickering, 1999; Riazantseva, 2001; Towell et al., 1996; Trofimovich & Baker, 2006). These nonnative pausing patterns correlate with lower fluency ratings and higher accentedness ratings (Bosker et al., 2013; Cucchiaroni et al., 2002; Derwing et al., 2004; Kahng, 2018; Kang, 2010; Trofimovich & Baker, 2006; Wennerstrom, 2000). Because using nonnative hesitation sounds is also a pausing pattern found in nonnative speech, and because the phonetic realization of the hesitation sound did not affect listener fluency judgments in Experiment 1A, it could be predicted that the phonetic realization of a hesitation sound would not affect accentedness ratings. However, it is also possible that compared to fluency judgments, accentedness judgments would be more influenced by the acoustic differences between native and nonnative hesitation sounds. In order to investigate whether the phonetic realization of a hesitation sound affects accentedness judgments, a second Likert scale rating experiment, Experiment 1B, was conducted.

### 6.1. EXPERIMENT 1B: PROCEDURE

The procedures for Experiment 1A and Experiment 1B are identical except for the directions that participants were given to rate sentences. In Experiment 1B, participants were given the following instructions about how to rate sentences for accentedness:

In this experiment, you will listen to English sentences and make judgments about how accented each sentence sounds. You will rate accentedness on a scale of 1-9.

It is important that you rate each sentence individually even if you think that you've heard the speaker's voice in a previous sentence.

1 = speaker does not at all have a nonnative accent in this sentence

9 = speaker has a very strong nonnative accent in this sentence

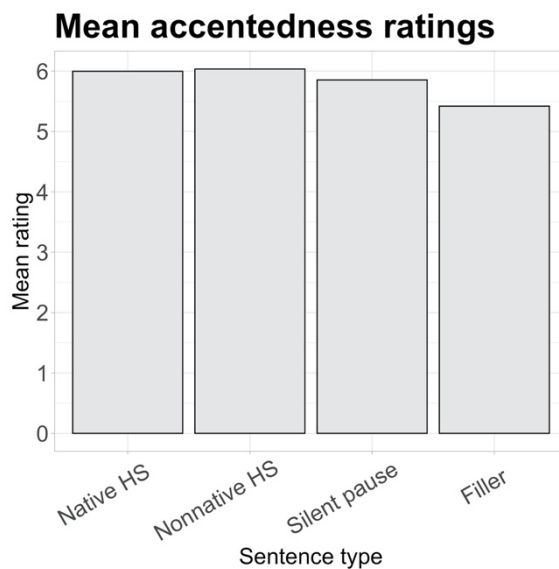
## 6.2. EXPERIMENT 1B: PARTICIPANTS

Fifty native English listeners participated in this experiment. When asked to self-report gender, 29 participants identified as female, 20 identified as male, and one participant identified as a transgender man. Participants had a mean age of 30.7 years, and the age range for participants was 18 to 65 years. Twenty-four participants were MTurk workers, and all other participants were University of Oregon students from the Psychology and Linguistics Human Subjects Pool.

## 6.3. EXPERIMENT 1B: RESULTS

As in Experiment 1A, data was analyzed in R (R Core Team, 2020) using a linear mixed models analysis. The phonetic realization of the hesitation sound (whether the hesitation sound was native or nonnative) did not impact listener judgments about accent and was not a significant predictor of model fit ( $\chi^2 = 0.0815$ ,  $p = 0.7753$ ). Additionally, whether the pause in the sentence was filled or unfilled did not impact listeners' accentedness ratings ( $\chi^2 = 1.0158$ ,  $p = 0.3135$ ). However, listeners judged stimuli sentences that contained a pause as more accented than filler sentences that did not ( $\chi^2 = 5.6127$ ,  $p = 0.01783$ ).

In summary, when participants listened to sentences and rated them for accentedness, neither the phonetic realization of the hesitation sound nor whether a pause was filled or unfilled impacted either fluency or accentedness ratings. However, whether a sentence contained a pause did impact listener ratings; sentences containing a pause of any kind were deemed more accented than sentences without pauses. Figure 2 shows the average accentedness ratings for sentences with each sentence type.



**Figure 2:** Mean accentedness ratings for Experiment 1B

## 7. EXPERIMENT 2A

The results of experiments 1A and 1B show that neither the phonetic realization of a hesitation sound nor whether a pause was filled or unfilled impacted listener judgments about accentedness or fluency. In order to investigate whether listener judgments would change if participants completed a different task to evaluate sentences for accentedness and fluency, two experiments with a forced choice task were designed and conducted after analyzing the results of Experiments 1A and 1B. In Experiments 2A and 2B, participants listened to two versions of the same sentence with different types of pauses and chose which sentence sounded more fluent or more accented.

### 7.1. EXPERIMENT 2A: PROCEDURE

As described above, the process for creating stimuli yielded three variations of the same sentence: one with an native (English) hesitation sound, one with a nonnative (Spanish) hesitation sound, and one with a silent pause. In Experiment 2A, participants listened to two variations of the same sentence (each sentence contained a different pause type) and chose which sentence made the speaker sound more fluent. Participants were given the following instructions:

In this experiment, you will listen to speech and make judgments about how fluent the speaker sounds in English in each sentence. Fluency means how well someone speaks a language.

You will listen to two versions of a sentence and decide which version makes the speaker sound more fluent in English.

During each trial, the participants were asked, “Does the speaker sound more fluent in English in sentence 1 or sentence 2?”

In total, participants heard 20 pairs of sentences that contained an English and a Spanish pause, 10 pairs of sentences that contained a silent and an English pause, and 10 pairs of sentences that contained a silent and a Spanish pause. No filler sentences were used in Experiments 2A or 2B.

### 7.2. EXPERIMENT 2A: PARTICIPANTS

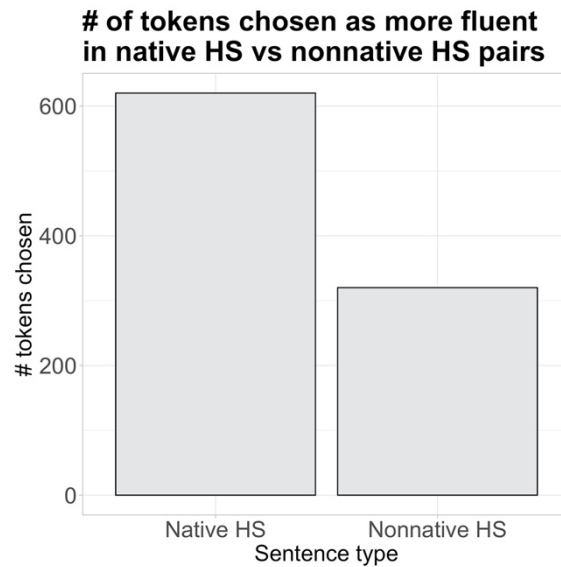
Forty-seven native English listeners participated in Experiment 2A. Thirty-seven participants were female, nine were male, and one participant identified as non-binary. Participants had a mean range of 23.3 years and the age range for participants was 18 to 67 years. Eleven participants were MTurk workers, and all other participants were University of Oregon students from the Psychology and Linguistics Human Subjects Pool.

### 7.3. EXPERIMENT 2A: RESULTS

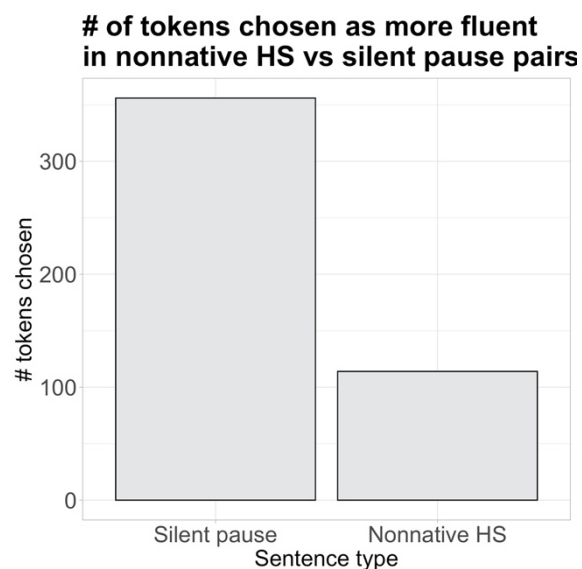
Data were analyzed using chi-squared tests in R (R Core Team, 2020). Listeners judged sentences with native (English) hesitation sounds as more fluent than sentences with nonnative (Spanish) hesitation sounds ( $\chi^2 = 95.745$ ,  $p < 0.001$ ). Participants also judged sentences with silent pauses as more fluent compared to sentences with Spanish hesitation sounds ( $\chi^2 = 124.6$ ,  $p < 0.001$ ). In addition, sentences with silent pauses were judged as more fluent than sentences with English hesitation sounds ( $\chi^2 = 17.234$ ,  $p < 0.001$ ).



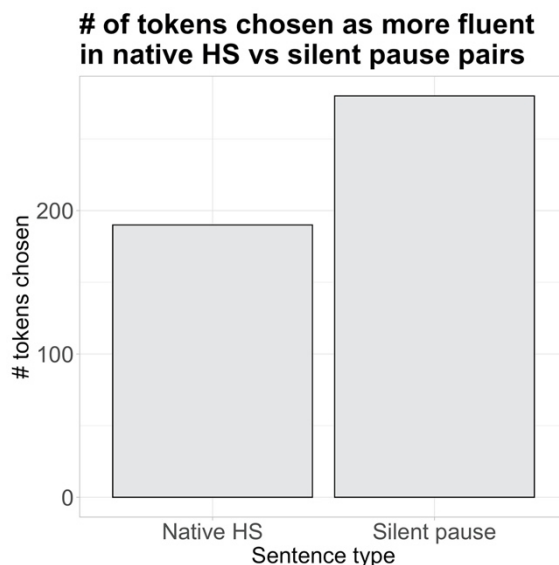
In other words, the results of Experiment 2A show that in a forced choice task where listeners were asked to evaluate sentences for fluency, the phonetic realization of hesitation sounds did influence listener judgments. Sentences with native (English) pauses were considered more fluent than sentences containing nonnative (Spanish) pauses. Additionally, sentences with silent pauses were judged as more fluent compared to sentences with nonnative hesitation sounds. Also, compared to sentences that contained native filled pauses, sentences with silent pauses were deemed more fluent. Figures 3 through 5 show the number of sentences with silent pauses, nonnative hesitation sounds, and native hesitation sounds that were chosen as more fluent in each pair type.



**Figure 3:** Number of sentences (tokens) selected as more fluent in native/nonnative hesitation sound pairs in Experiment 2A



**Figure 4:** Number of sentences (tokens) selected as more fluent in nonnative hesitation sound/silent pause pairs in Experiment 2A



**Figure 5:** Number of sentences (tokens) selected as more fluent in native hesitation sound/silent pause pairs in Experiment 2A

## 8. EXPERIMENT 2B

Experiment 2A shows that nonnative hesitation sounds decreased fluency ratings. To investigate whether the phonetic realization of a hesitation sound also affected perceptions about accentedness, a second forced choice task experiment, Experiment 2B, was also conducted.

### 8.1. EXPERIMENT 2B: PROCEDURE

The procedure for Experiment 2B was identical to Experiment 2A's procedure except that participants were instructed to choose which sentence sounded more accented. Participants received the following instructions:

In this experiment, you will listen to speech and make judgments about how accented each sentence sounds.

You will listen to two versions of a sentence and decide which version makes the speaker sound like they have a stronger nonnative accent.

During each trial, participants were asked, "Does it sound like the speaker has a stronger nonnative accent in sentence 1 or sentence 2?"

### 8.2. EXPERIMENT 2B: PARTICIPANTS

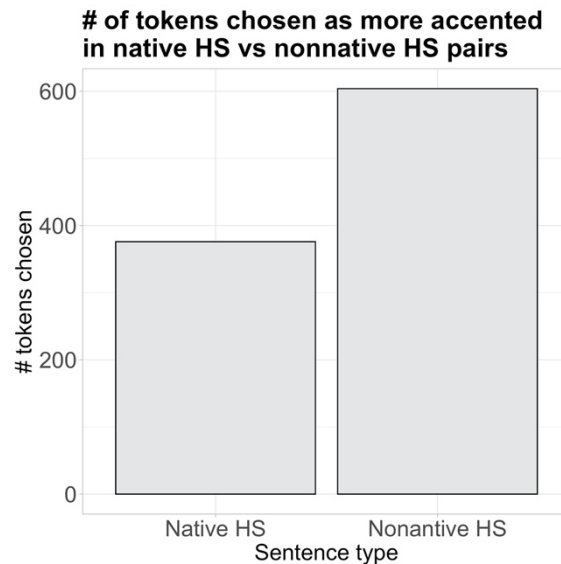
Forty-nine native English listeners participated in this experiment. Forty participants were female, eight were male, and one participant identified as non-binary. Participants had a mean age of 25.3 years and the age range for participants was 18 to 54 years. Sixteen participants were

MTurk workers, and all other participants were University of Oregon students from the Psychology and Linguistics Human Subjects Pool.

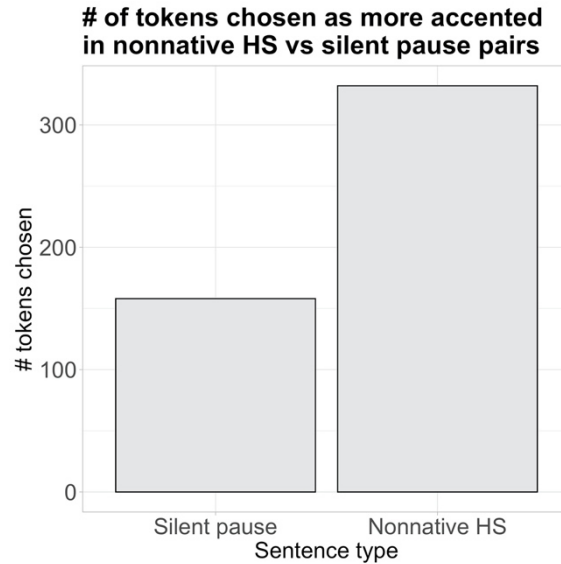
### 8.3. EXPERIMENT 2B: RESULTS

As in Experiment 2A, data was analyzed using chi-squared tests in R (R Core Team, 2020). Listeners judged sentences with nonnative (Spanish) hesitation sounds as more accented than sentences with native (English) hesitation sounds ( $\chi^2 = 53.045$ ,  $p < 0.001$ ). Listeners also deemed sentences with a nonnative (Spanish) hesitation sound to be more accented than sentences with a silent pause ( $\chi^2 = 61.788$ ,  $p < 0.001$ ). There was not a significant difference in accentedness judgments for sentences that contained an English hesitation sound versus a silent pause ( $\chi^2 = 2.3592$ ,  $p = 0.1245$ ).

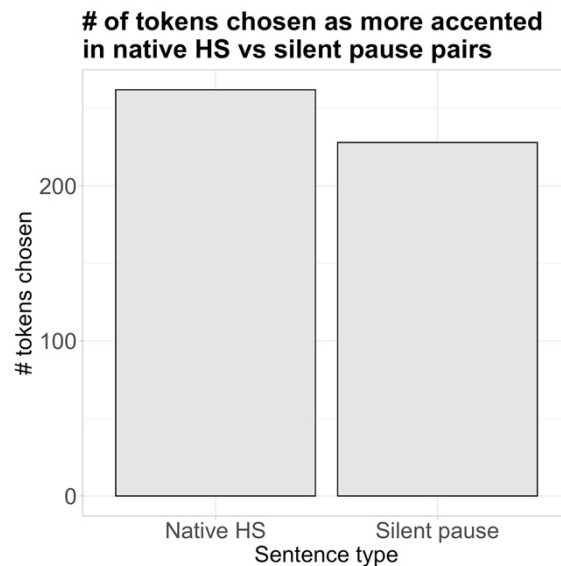
Data analysis reveals that when participants were asked to choose which sentence sounded more accented in a forced choice task, listeners judged sentences with nonnative (Spanish) hesitation sounds to be more accented compared to sentences with native (English) hesitation sounds. Listeners also judged sentences with nonnative (Spanish) filled pauses as more accented compared to sentences with silent pauses. However, listeners judged sentences with native hesitation sounds and silent pauses similarly. Figures 6 through 8 show the number of sentences with silent pauses, nonnative hesitation sounds, and native hesitation sounds chosen as more accented in each pair type.



**Figure 6:** Number of sentences (tokens) selected as more accented in native/nonnative hesitation sound pairs in Experiment 2B



**Figure 7:** Number of sentences (tokens) selected as more accented in nonnative hesitation sound/silent pause pairs in Experiment 2B



**Figure 8:** Number of sentences (tokens) selected as more accented in native hesitation sound/silent pause pairs in Experiment 2B

## 9. DISCUSSION

This study asks: Does the phonetic realization of the hesitation sound influence listener judgments about accentedness and fluency? In other words, how does using native versus nonnative hesitation sounds impact listener accentedness and fluency judgments? To answer these questions, four speech perception experiments were conducted. Two of the four

experiments, Experiments 1A and 1B, utilized a Likert scale rating task where listeners heard sentences with native (English), nonnative (Spanish), and silent pauses in addition to filler sentences that did not contain pauses. Listeners rated sentences for fluency and accentedness in Experiments 1A and 1B, respectively. The results of Experiments 1A and 1B reveal that the phonetic realization of the hesitation sound did not impact accentedness or fluency ratings.

However, in Experiments 2A and 2B, the phonetic realization of the hesitation sound did impact listener judgments about accentedness and fluency. In Experiments 2A and 2B, which utilized a forced choice task, participants heard two sentences that were identical except that they contained different types of pauses. Participants heard sentence pairs with a native and a nonnative hesitation sound, a native hesitation sound and a silent pause, and a nonnative hesitation sound and a silent pause, respectively. Listeners chose which sentence from each pair sounded more fluent in Experiment 2A and which sounded more accented in Experiment 2B. The results of Experiments 2A and 2B reveal that listeners judged sentences with nonnative hesitation sounds as less fluent and more accented than sentences with native hesitation sounds. The difference in results between Experiments 1A and 1B and Experiments 2A and 2B indicates that the different tasks used in these experiments affected how listeners made fluency and accent judgments.

One explanation for the difference in results between these experiments is that pauses were more salient to listeners in the forced choice task than they were in the Likert scale rating task. In each trial of the forced choice experiments, participants heard two sentences that were identical except that each sentence contained a different type of pause. That the two sentences heard in sequence only differed in the type of pause they contained perhaps made the pauses more prominent to listeners. Also, structural aspects of the Likert scale rating task in Experiments 1A and 1B could have decreased the relative salience of pauses and lessened the influence of pauses on listeners' ratings. During the Likert scale rating task, participants heard distinct sentences in a randomized order. Unlike Experiment 2A and 2B participants, Experiment 1A and 1B participants never heard the same sentence more than once, let alone the same sentence consecutively with a different pause. Also, listeners who completed the Likert scale rating task heard filler sentences that did not contain pauses in addition to those that did contain pauses. In contrast, forced choice task participants only heard sentences with pauses. Increased pause salience in the forced choice experiments could have made the type of pause more influential in listener judgments; decreased pause salience in the Likert scale rating task could have made the pause less influential.

Another explanation for the contrasting results in Experiments 1A and 1B and Experiments 2A and 2B is that listeners made more holistic judgments in the Likert scale rating task and more fine-grained judgments in the forced choice task. That is, listeners in the Likert scale rating task may have used top-down cues that caused pauses to have less influence on their judgments. In contrast, forced choice listeners may have made more bottom-up judgments that caused pauses to have more influence. In other words, in the Likert scale rating task, participants perhaps made holistic decisions about fluency and accentedness based on their general impression of the sentence. Conversely, forced choice participants may have examined individual parts of each speech signal, identified what differentiated them, and based their judgment on the differentiating feature.

The results of this study suggest that when listeners make holistic judgments about nonnative speech — or in situations where nonnative speech is not scrutinized — using nonnative hesitation sounds is unlikely to negatively impact listener opinions about a nonnative speaker's accent or how well they speak a language. This finding has important implications for how learners treat pausing when practicing their second language and for how nonnative speech is perceived in real-world situations. Research shows that practicing speaking a second language is one of the best ways to increase proficiency (Izumi et al., 1999; Toth, 2006; Valezy & Spada, 2006). Yet research also shows that speakers are more likely to pause when speaking their second language (Pickering, 1999; Riazantseva, 2001). As learners' proficiency improves, it is likely that their pausing patterns will come to resemble those of native speakers (Erker & Bruso, 2017). This study suggests that learners should pause however it feels natural to them, especially in everyday situations where speech is unlikely to be closely examined. However, this study also suggests that in situations where listeners judge nonnative speech in a fine-grained way, using nonnative hesitation sounds may lead to negative perceptions about a speaker's accent and fluency. Therefore, education about the phonetic quality of hesitation sounds may be a useful addition to language teaching curriculum.

This study demonstrates the need to reduce stigma around pauses, especially regarding nonnative speech. This study also has important implications for future work investigating perception of nonnative speech. Results show that both the context in which listeners hear nonnative speech and the task that listeners perform to evaluate speech affect listener judgments. The effect of task on the perceptual consequences of filled and unfilled pauses should be further investigated. Additionally, the results of this study add complexity to speech perception literature suggesting that filled and unfilled (silent) pauses have different perceptual consequences. For example, it has been found that unfilled pauses negatively affect listener perceptions about accentedness and comprehensibility (Kang, 2010). However, the results of Experiments 1A and 1B showed that whether a pause is filled or unfilled did not affect listener judgments about accent or fluency. Yet, Experiment 2A and 2B results showed that sentences with native (English) filled pauses were judged as less fluent and more accented than silent pauses. More research is needed to fully understand the perceptual effects of filled versus unfilled pauses in all speech. Additionally, because this project was limited to the perception of male speech, future studies should investigate the effect of hesitation sound phonetic quality on the perception of non-male speech.

## 10. CONCLUSION

This study examined the effect of hesitation sound phonetic quality on perception of language fluency and accent. In two of the four experiments conducted (Experiments 1A and 1B), the phonetic realization of hesitation sounds did not impact listener judgments about accent or fluency. However, in the other two experiments conducted (Experiments 2A and 2B), which utilized a different experimental design, sentences with nonnative hesitation sounds were judged as less fluent and more accented than sentences with native hesitation sounds. Two explanations are offered for the differing results between experiments. It is possible that pauses were perceived as more salient in Experiments 2A and 2B than they were in Experiments 1A and 1B and the salience of pauses affected listener judgments. It is also possible that listeners made fine-grained

or holistic judgments depending on the task they performed to evaluate speech. The results of this study suggest that the type of task that listeners perform to evaluate speech affects listener judgments about fluency and accentedness. With some tasks, nonnative hesitation sounds can cause speech to be perceived as less fluent and more accented. The results of this study also show that the phonetic realization of a hesitation sound can have significant perceptual consequences.

## ACKNOWLEDGEMENTS

I am profoundly grateful to my mentor and brilliant advisor Dr. Melissa Baese-Berk for her unfailing support, encouragement, and guidance during this project and throughout my undergraduate experience. Thank you to the vibrant community at Dr. Baese-Berk's Speech Perception and Production Lab for your collective wisdom. Finally and importantly, thank you to the Office of the Vice President for Research and Innovation, the Acoustical Society of America, and the Center for Undergraduate Research and Engagement for funding this research.

## REFERENCES

- Anderson-Hsieh, J., & Venkatagiri, H. (1994). Syllable Duration and Pausing in the Speech of Chinese ESL Speakers. *TESOL Quarterly*, 28(4), 807–812. <https://doi.org/10.2307/3587566>
- Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* (6.1.13) [Computer software]. <http://www.praat.org/>
- Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30(2), 159–175. <https://doi.org/10.1177/0265532212455394>
- Boyle, J. (2015). *Vocal pitch manipulation and credibility in the courtroom: Differences between male and female expert witnesses*. <https://www.proquest.com/openview/791f0467e9d4a737e286609c32cb037a/1?pq-origsite=gscholar&cbl=18750>
- Bradlow, A. (n.d.). *ALLSTAR: Archive of L1 and L2 Scripted and Spontaneous Transcripts And Recordings*. Retrieved July 21, 2020, from [https://groups.linguistics.northwestern.edu/speech\\_comm\\_group/allstar2/#!/](https://groups.linguistics.northwestern.edu/speech_comm_group/allstar2/#!/)
- Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America*, 97(3), 1916–1924. <https://doi.org/10.1121/1.412064>
- Brennan, S. E., & Williams, M. (1995). The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *Journal of Memory and Language*, 34(3), 383–398. <https://doi.org/10.1006/jmla.1995.1017>
- Campillos-Llanos, L., & Plá, M. A. (2009). *Speech Disfluencies in Formal Context. Analysis based on Spontaneous Speech Corpora*.
- Cenoz, J. (1998). *Pauses and Communication Strategies in Second Language Speech*.

- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535–544.  
[https://doi.org/10.1016/S0346-251X\(97\)00046-8](https://doi.org/10.1016/S0346-251X(97)00046-8)
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111. [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3)
- Cohen, S. D. (2012, October 22). *Tips on Public Speaking: Eliminating the Dreaded “Um.”*  
<https://blog.dce.harvard.edu/extension/tips-public-speaking-eliminating-dreaded-um>
- Cucchiari, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, 111, 2862–2873. <https://doi.org/10.1121/1.1471894>
- Davies, A. (2003). *The Native Speaker: Myth and Reality*. Multilingual Matters.
- Davies, A., & Elder, C. (2008). *The Handbook of Applied Linguistics*. John Wiley & Sons.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, 54(4), 655–679.  
<https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- Erbaugh, M. (1987). A uniform pause and error strategy for native and non-native speakers. In R. S. Tomlin (Ed.), *Coherence and grounding in discourse: Outcome of a symposium, Eugene, Oregon, June* (pp. 110–130). Amsterdam ; Philadelphia : J. Benjamins Pub. Co.  
<https://books.google.com/books?hl=en&lr=&id=v8FOAgAAQBAJ&oi=fnd&pg=PA109&dq=unfilled+vs+filled+pauses+nonnative&ots=3LYeYnnyUu&sig=ko09RECbDtMJ3bC1JjooQ91AfVE#v=onepage&q=unfilled%20vs%20filled%20pauses%20onnonnative&f=false>
- Erker, D., & Brusco, J. (2017). Uh, bueno, em ...: Filled pauses as a site of contact-induced change in Boston Spanish. *Language Variation and Change*, 29(2), 205–244.  
<https://doi.org/10.1017/S0954394517000102>
- Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, 29(2), 320–326. <https://doi.org/10.3758/bf03194926>
- Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*, 65(2), 161–175.  
<https://doi.org/10.1016/j.jml.2011.03.004>
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. Academic P.
- Hlavac, J. (2011). Hesitation and monitoring phenomena in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation? *Journal of Pragmatics*, 43(15), 3793–3806. <https://doi.org/10.1016/j.pragma.2011.09.008>
- Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed Levels of Second Language Speaking Proficiency: How Distinct? *Applied Linguistics*, 29(1), 24–49.  
<https://doi.org/10.1093/applin/amm017>



- Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421–452.
- Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, 39(3), 569–591. <https://doi.org/10.1017/S0142716417000534>
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38(2), 301–315. <https://doi.org/10.1016/j.system.2010.01.005>
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698–2704. <https://doi.org/10.1098/rspb.2012.0311>
- Kormos, J., & Dénes, Mariann. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164. <https://doi.org/10.1016/j.system.2004.01.001>
- Lennon, P. (1990). Investigating Fluency in EFL: A Quantitative Approach\*. *Language Learning*, 40(3), 387–417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- Maclay, H., & Osgood, C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. *WORD*, 15(1), 19–44. <https://doi.org/10.1080/00437956.1959.11659682>
- Pickering, L. (1999). *An analysis of prosodic systems in the classroom discourse of native speaker and nonnative speaker teaching assistants*. <https://ufdc.ufl.edu/AA00014294/00001>
- Qualtrics (July 2020). (2005). [Computer software]. Qualtrics.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition*, 23, 497–526. <https://doi.org/10.1017/S027226310100403X>
- Riegel, D. G. (2018, August 21). *Your, Um, Filler Words Are, Like, Killing Your Credibility. Here's How to, Uh, Fix This in 4 Weeks*. Inc.Com. <https://www.inc.com/deborah-grayson-riegel/how-to-get-rid-of-filler-words-sound-more-professional-in-4-weeks.html>
- Roggia, A. B. (2012). Eh as a polyfunctional discourse marker in Dominican Spanish. *Journal of Pragmatics*, 44(13), 1783–1798. <https://doi.org/10.1016/j.pragma.2012.08.010>
- Rose, R. L. (2017). A Comparison of Form and Temporal Characteristics of Filled Pauses in L1 Japanese and L2 English. *Journal of the Phonetic Society of Japan*, 21(3), 33–40. [https://doi.org/10.24467/onseikenkyu.21.3\\_33](https://doi.org/10.24467/onseikenkyu.21.3_33)

- Toth, P. D. (2006). Processing Instruction and a Role for Output in Second Language Acquisition. *Language Learning*, 56(2), 319–385. <https://doi.org/10.1111/j.0023-8333.2006.00349.x>
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The Development of Fluency in Advanced Learners of French. *Applied Linguistics*, 17(1), 84–119. <https://doi.org/10.1093/applin/17.1.84>
- Trofimovich, P., & Baker, W. (2006). LEARNING SECOND LANGUAGE SUPRASEGMENTALS: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquisition*, 28(1), 1–30. <https://doi.org/10.1017/S0272263106060013>
- Valezy, J. R., & Spada, N. (2006, May 9). 4. *The effectiveness of corrective feedback for the acquisition of L2 grammar*. John Benjamins Publishing Company. <https://benjamins.com/catalog/llt.13.09val>
- Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication*, 50(2), 81–94. <https://doi.org/10.1016/j.specom.2007.06.002>
- Wennerstrom, A. (2000). The role of intonation in second language fluency. *Perspectives on Fluency*, 2000, ISBN 0-472-11028-4, Págs. 102-127, 102–127. <https://dialnet.unirioja.es/servlet/articulo?codigo=5732207>