

Chapter

An Acoustic Study on the Use of Fillers in Spanish as a Foreign Language Acquisition

María J. Machuca

Abstract

Filled pauses are a vital component of foreign language learners' communicative competence. Both instructors and students should be cognizant of its importance and employ various communication techniques to reduce the foreign accent. The most common Spanish filler is /e/. This study aims to investigate the vocalic fillers used by SFL learners. Twenty-four speakers with different L1s (English, French, and Russian) and language proficiency (intermediate and advanced) participated in the experiment. As these languages use distinct vocalic elements to fill pauses, the linguistic transfer may occur. Participants engaged in two semi-spontaneous tasks for data elicitation. Their fillers were categorized as either /e/ or non-/e/. Additionally, F1 and F2 values for filler classified as /e/ were compared between the control and experimental groups. In terms of fillers, the results indicate a linguistic transfer from the learners' L1 to Spanish. Also, no difference was found for the F1 and F2 of /e/ between SFL learners and native Spanish speakers. In addition, learners with advanced proficiency were more likely to yield the correct production than those with intermediate proficiency.

Keywords: fillers, pauses, proficiency, foreign language acquisition, fillers segmentation

1. Introduction

Hesitations are a common communicative strategy in oral discourse across all languages. In fact, speakers use them to plan what they will say during a spontaneous speech that was not previously prepared. This singularity of not preparing what will be said in advance causes speakers who are preparing their speech in real time to correct what they are saying or to introduce hesitations in order to organize the oral language they are using while speaking. Speakers require time to develop this cognitive process; hesitations give them this time. Some hesitations are unique to each language [1], while others distinguish among distinct linguistic varieties [2]. For example, European Spanish speakers typically do not employ the same hesitations as their American counterparts [3]. In Argentina and Venezuela, “este” (this) is used as a lexical filler, whereas in Guatemala and Spain, “pues” (then) is more common [4, 5].

Maclay [6] classified four types of hesitations in English that had no semantic relevance in the discourse: repetitions that are not used to emphasize, false starts without or with self-correction, filled pauses (including vocalic or consonant lengthening), and silences. However, some authors distinguish between lengthening and filled pauses since lengthening is semantically closer to discourse than filled pauses [5]. Furthermore, although pauses and lengthening serve the same purpose in spontaneous speech, they exhibit some differences. Duration in Spanish can differentiate between both hesitations. According to Machuca [7], lengthening can achieve a significantly longer duration (2667 ms) than filled pauses (1292 ms). Blondet [5] demonstrates that Spanish lengthening is, on average, longer than pauses (637 ms vs. 414 ms). In this study, we will only examine non-lexical filled pauses, that is, those vocalizations that are not part of lexical components of the language.

Filled pauses are claimed to be one of the most widely used linguistic strategies (cf. [8, 9]), as it not only serves to indicate the appropriate time to take a turn in conversation [10, 11] but also provides time for speakers to plan, reorganize, and execute their oral discourse ([12, 13], among others). In addition, Goldman-Eisler [14] noted that the words following this type of hesitation might contain pertinent information. They are used in cases where there is an interaction among conversational participants, as well as when the participants know they will not be interrupted during their speech [7]. Thus, it is a natural phenomenon in speech, even if the speaker is unaware of it and the listener does not appear to perceive it [15, 16]. According to Machuca [7], a filled pause placed in the middle of a sentence with a duration of 156 ms is not perceived by the listeners because the meaning of the interaction does not change since the filled pauses are not associated with the specified lexical unit.

Due to the fact that they are natural phenomena and frequently appear in speech, they have been analyzed to explain certain actions in various disciplines, such as speech technology, speaker identification in forensic phonetics, categorization of different frontotemporal diseases that affect speech, and a criterion for defining fluency in a foreign language learning, among others.

In speech technologies, two different perspectives exist. On the one hand, fillers are extracted from the systems, reducing the accuracy of automatic speech recognition (ASR) systems. A system that automatically recognizes filled pauses will significantly reduce errors in recognizing spontaneous speech, as hesitation is the most frequent phenomenon across languages [17–19]. On the other hand, fillers are incorporated into the systems. By incorporating fillers into the design of conversational robots, their speech becomes more natural and human-like, and as reported by some authors, their perception of turn-taking speed is altered [20].

In forensic applications, filled pauses are utilized as well. In conjunction with other hesitation phenomena, this type of pause helps to identify a speaker [21, 22]. However, if we focus solely on the filled pause, Braun [23] concluded that the proportion of segments with which a pause is filled varies among speakers. McDougall [24] obtained similar results: the selection of filled pauses is influenced by individual differences. Consequently, the acoustic features of vocalization can provide information about the speaker, as he selects one of the phonic emissions and frequently uses filled pauses in his speech [25–27].

In the acquisition of a foreign language, filled pauses are also used, even though CEFR [28] does not adequately account for these phenomena in the students' assessment, despite being considered one of the most accurate ways to determine whether a foreign language is being acquired successfully [5]. Aside from whether vowels that

are used as fillers are included in the phonological system of a given language since there is no consensus on this issue [29], the focus of this study is on whether students of foreign languages use vocalic filled pauses of the language they are learning or whether they continue to use the filler of their native language, even at advanced proficiency levels. This paper analyzes the filled pauses of native English, Russian, and French speakers with a high level of Spanish. As previously demonstrated by Candea [1], native speakers of these three languages tend to use different fillers compared with Spanish speakers. These authors compared the vocalic segments of eight languages and concluded that frequency values corresponding to central position do not appear to be universal, despite being the most common timbre for the majority of the languages considered in that study.

The two most common fillers in English are *um* and *uh*, which are pronounced as schwa with or without nasal consonant ending. They display varying proportions based on the English Variety [30]. Hughes [31], in their study analyzing fillers in Southern British English, also indicate a considerable variation among speakers (both within and between them), but F1 and F2 values would correspond to schwa. F1 values are detected between 450 Hz and 700 Hz, while F2 values are detected between 1250 Hz and 1550 Hz. Comparing the frequency trajectories of the two formants revealed similar values for each speaker (p. 22).

In French, the filler represented by *euh* corresponds to the central vowel's mean value. F1 frequencies are approximately 470 Hz for men and 523 Hz for women, while F2 frequencies are 1464 Hz for men and 1659 Hz for women [1]. The authors compared the timbre of these fillers to that of similar intralexical vocalic segments. They found that the F1 values of the fillers were slightly higher, possibly because the intralexical vocalic segments are considerably shorter than the fillers. F1 frequency values are the same when comparing speakers of different ages, whereas F2 values do not exhibit the same behavior [32].

In terms of Russian, Stepanova [13] compared the spectral characteristics of speakers' vocalizations of hesitation pauses with stressed vowels /a/ and /e/ within words to determine if the filler had acoustic characteristics similar to one of these vowels. She observed that F1 values in fillers are realized as a vowel between /a/ and /e/, whereas F2 frequencies have the same value as /a/. In Gil [33], 660 filler vowels produced by Russian speakers were analyzed and compared with Russian /a, e/. The average frequency of F1 was 519 Hz, and the frequency of F2 was 1222 Hz. F1 values in fillers indicate intermediate frequencies between the two vowels analyzed, while F2 values are slightly lower than /a/.

In Spanish, according to Machuca [7], there are three vocalizations for filling hesitations. In their study, 61.5% of the fillers correspond to the vowel [e], the most frequent filler, 32% correspond to a nasal murmur, and 7% include vocalizations that could be transcribed as [a] and other fillers unable to be categorized in any of the previous cases (p. 86). In their later study [34], based on the same corpus of spontaneous speech, analyzed the fillers categorized as [e] in male speakers and compared them with the same vowel in both stressed and unstressed positions. The F1 values were similar across the three segments analyzed, while the F2 values were consistently higher in the vocalic segment corresponding to fillers, 463 Hz for the F1 values and 1800 Hz for the F2 values. Villa [35] also analyzed the vocalic fillers in Spanish. However, they compared their acoustic characteristics to the /e/ vowel produced by the same speakers in a reading task. In F1, the results are very similar (473 for the filler and 467 for the vowel /e/), but in F2, the filler has higher frequencies than the vowel (1903 Hz versus 1707 Hz).

In foreign language learning, non-native speakers tend to use more unfilled pauses than fillers [36]. In contrast, advanced learners tend to use fillers with different vocalizations in order to sound like native speakers [37]. But the instruction of foreign language fillers is “a neglected aspect of teaching L2” [38]. Actually, these phenomena are not mentioned in [28], with the exception of the term fluency, which is defined as “the ability to deploy the resources in real time to produce connected discourse with normal rhythm and intonation, free from hesitations, false starts, etc.” (p. 16). Learners frequently use the filled pauses of their native language when they produce speech hesitation phenomena [39]. Because learners may have difficulty perceiving the phonetic difference among these sounds in two languages, teachers should emphasize these differences so that learners begin to vocalize fillers as vocalic segments in a foreign language. Consequently, the foreign accent will be perceived less strongly in the learners.

This study aims to analyze the vocalic filler of Spanish learners whose native languages have vocalic segments for fillers that are vastly distinct from those of the foreign language they are learning, Spanish. The research assumes that the vocalic element for the filler in Spanish, as we have seen, is similar to an [e], which differs from English and French with a vocalic segment in fillers tending to the values of a schwa and the Russian, which fillers show values of F2 similar to the vowel /a/. In addition, we will consider different levels of proficiency to determine if they have already replaced advanced-level filled pauses in their native language with a vocalic element used in the foreign language they are learning.

2. Experimental design

2.1 Corpus and speakers

For this analysis, we utilized the EMULANDO corpus [40]. The main objective of this corpus was to compare natural and masked foreign accents in Spanish for forensic phonetics. The fillers were extracted from two semi-spontaneous tasks. Speakers with intermediate and advanced proficiency levels were chosen, as they were required to participate in two different scenarios using spontaneous speech from a storyboard containing data. In the first task, speakers were required to inquire about studying in Spain. The second task required the speakers to use embarrassing photographs of a public figure to extort money from him. Participants were native speakers of English, French, and Russian, and they all had to perform these tasks in Spanish. A control group comprising native Spanish speakers was also included.

In total, thirty-two speakers have been analyzed, including eight native Spanish speakers as a control group, three English speakers with high Spanish proficiency and four with intermediate proficiency, five French speakers with advanced Spanish proficiency and four with intermediate proficiency, and four Russian speakers with advanced Spanish proficiency and four with intermediate proficiency. A total of 844 cases have been analyzed, 420 of which involved filled pauses produced by a vocalic segment similar to /e/ and 424 of which involved filled pauses produced by a vowel other than /e/. Native speakers of English produced 115 instances of vocalic fillers, French speakers 276 instances, Russian speakers 268 instances, and Spanish speakers 185 instances of /e/-sounding fillers.

2.2 Fillers segmentation and annotation

Fillers were segmented with Praat [41] by four Spanish-native annotators. In the samples of non-native speakers, the annotators were required to indicate whether the vocalic element filler was perceived as /e/, the most frequent filler in Spanish. If it was not perceived as /e/, it was classified as a vowel that did not sound like /e/. This study analyzed segments labeled as /e/ by four annotators prior to data extraction. If there was no consensus among the four annotators, the filler was recategorized as non-/e/. Then, we extracted mid-point formant values (F1 and F2) from those categorized as /e/ and those not classified as /e/. The values of F1 and F2 in the native samples served as a comparison baseline.

2.3 Statistical analysis

Based on the research questions that we have formulated, we have employed a variety of statistical tests according to the variables analyzed. The research questions of this study are as follows:

Question 1: Are the acoustic parameters of filled pauses perceptually classified as /e/ distinct from those classified as non-/e/?

In this case, a t-test is performed. Data is subsetted by gender, native language, and two fillers (/e/ and non-/e/). F1 and F2 are the dependent variables, while /e/ and non-/e/ categorization are independent variables.

Question 2: Does the filler categorized as /e/ vary depending on the native language?

Speaker was modeled as a random effect using Linear Mixed Effects Regression Model [42]. F1, F2 values as the dependent variables; filler (/e/ and non-/e/), L1 (English, French, and Russian), gender (male and female) are the independent variables.

Question 3: What frequency region does non-/e/ production fall under?

This question is addressed by comparing the distribution of F1 and F2 values in each native language group with the Spanish filler /e/.

Question 4: Does the proficiency of the speaker influence the selection of these filled pauses?

To answer this question, a Chi-square test was conducted with proficiency level as the independent variable and the vocalic segment used to fill the pause (/e/ and non-/e/) as the dependent variable.

3. Results

This section¹ presents results in the same order as the questions in Section 2. **Table 1** presents some descriptive information on the F1 and F2 values according to speakers' L1, filler, and sex. No cases of non /e/ for Russian male speakers have been found in our corpus.

Question 1: Are the acoustic parameters of filled pauses perceptually categorized as a /e/ different from those categorized as a non-/e/?

¹ The statistical treatment of the data would not have been possible without the help of Wenhao Li, Student of M.S. in Statistics at New York University.

| L1 | Filler | Sex | F1 | | F2 | |
|---------|--------|--------|--------|--------|---------|--------|
| | | | Mean | SD | Mean | SD |
| English | e | Female | 551.11 | 52.83 | 2290.31 | 149.99 |
| | | Male | 430.17 | 41.78 | 2122.82 | 131.54 |
| | non_e | Female | 795.00 | 122.58 | 1589.85 | 217.04 |
| | | Male | 520.53 | 107.90 | 1321.14 | 189.43 |
| French | e | Female | 570.19 | 90.18 | 2047.83 | 212.00 |
| | | Male | 431.82 | 39.74 | 1932.91 | 89.39 |
| | non_e | Female | 619.52 | 87.72 | 1632.85 | 163.70 |
| | | Male | 494.72 | 81.75 | 1578.23 | 138.34 |
| Russian | e | Female | 512.64 | 66.55 | 1982.77 | 211.24 |
| | | Male | 431.81 | 42.32 | 2086.24 | 169.48 |
| | non_e | Female | 760.64 | 137.53 | 1503.39 | 176.88 |
| | | Male | 529.00 | 90.76 | 1859.92 | 314.68 |
| Spanish | e | Female | 529.00 | 90.76 | 1859.92 | 314.68 |
| | | Male | 480.46 | 55.03 | 2017.37 | 196.52 |

Table 1. Mean and standard deviation (SD) of F1 (Hz) and F2 (Hz) values per language, filler, and sex.

| L1 | Sex | Parameter | Results |
|---------|--------|--------------|---|
| English | Female | F1 | $t_{\text{Welch}}(47.16) = -10.00, p = 3.09e-13, g_{\text{Hedges}} = -2.54, CI_{95\%} [-3.25, -1.82]$ |
| | | (N = 53) F2 | $t_{\text{Welch}}(50.01) = 13.87, p = 9.23e-19, g_{\text{Hedges}} = 3.70, CI_{95\%} [2.80, 4.58]$ |
| | Male | F1 | $t_{\text{Welch}}(53.77) = -4.67, p = 2.05e-05, g_{\text{Hedges}} = -1.09, CI_{95\%} [-1.59, -0.58]$ |
| | | (N = 62) F2 | $t_{\text{Welch}}(58.27) = 19.60, p = 2.48e-27, g_{\text{Hedges}} = 4.85, CI_{95\%} [3.84, 5.85]$ |
| French | Female | F1 | $t_{\text{Welch}}(191.96) = -3.88, p = 1.41e-04, g_{\text{Hedges}} = -0.55, CI_{95\%} [-0.84, -0.27]$ |
| | | (N = 197) F2 | $t_{\text{Welch}}(192.93) = 15.48, p = 1.14e-35, g_{\text{Hedges}} = 2.18, CI_{95\%} [1.83, 2.53]$ |
| | Male | F1 | $t_{\text{Welch}}(46.67) = -4.18, p = 1.28e-04, g_{\text{Hedges}} = -0.96, CI_{95\%} [-1.45, -0.47]$ |
| | | (N = 79) F2 | $t_{\text{Welch}}(55.5) = 13.14, p = 1.11e-18, g_{\text{Hedges}} = 3.00, CI_{95\%} [2.28, 3.71]$ |
| Russian | Female | F1 | $t_{\text{Welch}}(50.82) = -15.35, p = 9.97e-21, g_{\text{Hedges}} = -2.26, CI_{95\%} [-2.78, -1.73]$ |
| | | (N = 251) F2 | $t_{\text{Welch}}(27.85) = 10.93, p = 1.39e-11, g_{\text{Hedges}} = 2.39, CI_{95\%} [1.63, 3.15]$ |

Table 2. Significance level considering L1 and gender in /e/ and non /e/ fillers. The frequency dispersion area of the first two formants for all groups of learners is displayed in **Figure 1**.

Taking into account the two vocalic elements used to fill the pause, it could be stated that the differences are significant for each group with different native languages and, within each group, for both men and women. **Table 2** displays the statistical results in detail. As we can see, the comparison was not possible among

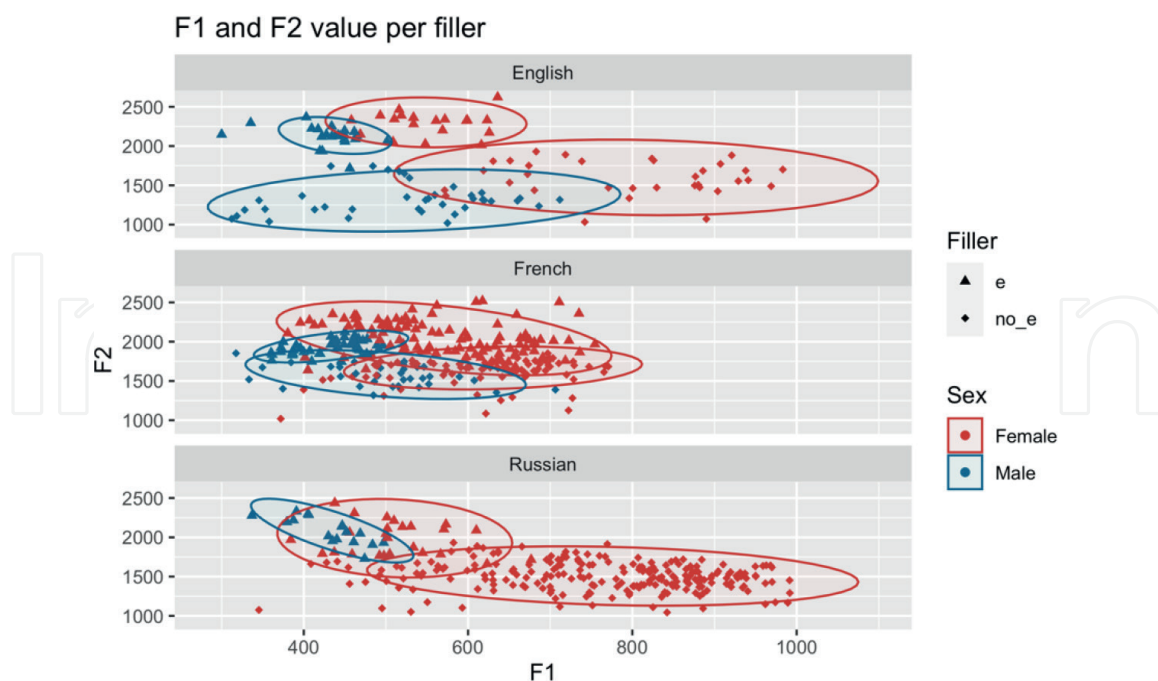


Figure 1.
Dispersion area of the first two formants for all learner groups. Fillers are distinguished by shape and sex by color.

male Russian speakers because all cases were classified as /e/. As later we will see in Question 4, these results are related to that male speakers have advanced Spanish proficiency.

Because the value of F1 is lower (less oral opening than the non-/e/ filler) and the value of F2 is higher (vowel located in a more frontal region than the non-/e/ filler), the fillers classified as /e/ are located in a frequency region to the left of the figure.

Question 2: Does the filler categorized as /e/ vary depending on the native language?

We have fitted four different linear mixed models. All models included the speaker as random effect. Results can be observed in **Table 3**.

Comparing the F1 values of filled pauses categorized by /e/ between learners and native Spanish speakers reveals no differences for women, but differences are observed for men.

The first model predicts F1 in female speakers with L1. The model's intercept is at 532.73. Within this model, no group differs significantly. English and French speakers present higher values, and F1 values for Russian speakers are lower. In terms of F1 values in men, the model's intercept is at 483.69. All groups are statistically significant and negative.

F2 values matching /e/ in native Spanish speakers to fillers categorized as /e/ in learners do not differ significantly, except for the group of female English speakers.

For F2 values in women, the model's intercept is at 1887.95. Within this model, the effect of L1 English is statistically significant and positive. Both French and Russian groups are statistically nonsignificant and positive. For F2 values in men, the model's intercept is at 1991.93. Within this model, the English and Russian groups are statistically nonsignificant and positive, while the French group is statistically nonsignificant but negative.

Figure 2 depicts all instances of /e/ and fillers used by English, French, Russian, and Spanish speakers. F1 values are higher in the French group than in the native

| Model | Predictor | β | β CI _{95%} | Results | Std. β | Std. β CI _{95%} |
|-----------|------------|---------|---------------------------|-----------------------------------|--------------|--------------------------------|
| F1 Female | L1 English | 18.37 | [-87.46, 124.21] | $t(179) = 0.34$, $p = 0.732$ | 0.21 | [-1.01, 1.44] |
| | L1 French | 40.76 | [-29.10, 110.63] | $t(179) = 1.15$, $p = 0.251$ | 0.47 | [-0.34, 1.28] |
| | L1 Russian | -5.84 | [-82.65, 70.96] | $t(179) = -0.15$, $p = 0.881$ | -0.07 | [-0.96, 0.82] |
| F2 Female | L1 English | 402.36 | [92.56, 712.15] | $t(179) = 2.56$, $p = 0.011$ | 1.58 | [0.36, 2.80] |
| | L1 French | 149.83 | [-54.03, 353.68] | $t(179) = 1.45$, $p = 0.149$ | 0.59 | [-0.21, 1.39] |
| | L1 Russian | 91.42 | [-131.72, 314.55] | $t(179) = 0.81$, $p = 0.420$ | 0.36 | [-0.52, 1.24] |
| F1 Male | L1 English | -53.52 | [-94.30, -12.73] | $t(229) = -2.59$, $p = 0.010$ | -0.97 | [-1.70, -0.23] |
| | L1 French | -50.72 | [-79.83, -21.62] | $t(229) = -3.43$, $p < .001$ | -0.91 | [-1.44, -0.39] |
| | L1 Russian | -51.88 | [-94.29, -9.46] | $t(229) = -2.41$, $p = 0.017$ | -0.94 | [-1.70, -0.17] |
| F2 Male | L1 English | 130.89 | [-70.84, 332.61] | $t(229) = 1.28$, $p = 0.202$ | 0.72 | [-0.39, 1.84] |
| | L1 French | -59.14 | [-199.64, 81.35] | $t(229) = -0.83$, $p = 0.408$ | -0.33 | [-1.11, 0.45] |
| | L1 Russian | 94.30 | [-110.59, 299.20] | $t(229) = 0.91$, $p = 0.365$ | 0.52 | [-0.61, 1.66] |

Table 3. Summarizes the results. All model's intercept corresponds to L1 = Spanish.

speaker group. F2 values are greater in the English group than in the native group. When compared with the native group, all learner groups exhibit greater variability.

Question 3: What is the frequency region for cases categorized as non /e/?

To determine the formant values' frequencies, we examined the F1 and F2 frequencies of all non-native speakers' fillers. **Figure 3** only displays filler classified as non-/e/, and Spanish /e/ has been included in the graph to provide a reference point. Based on the dispersion area of F1 and F2 values, we can see that the filler of Spanish speakers is relatively stable, unlike a schwa.

Generally speaking, the F1 and F2 values of the schwa are 500 Hz and 1500 Hz, respectively. Taking this into account, **Figure 3** demonstrates that French speakers tend to produce more schwa-like sounds than the other group. Russian speakers, on the other hand, produce a vowel with an F1 value of approximately 800 Hz that is more open, and their F2 value is lower than their schwa value. In contrast, the frequency areas of English speakers are quite dispersed, making it difficult to identify which vowel was produced. If we observe the location of /e/ in native speakers, we can conclude that learners employ the vocalic element they employ in their native language.

Question 4: Does the speaker's proficiency have an effect on the performance of these filled pauses?

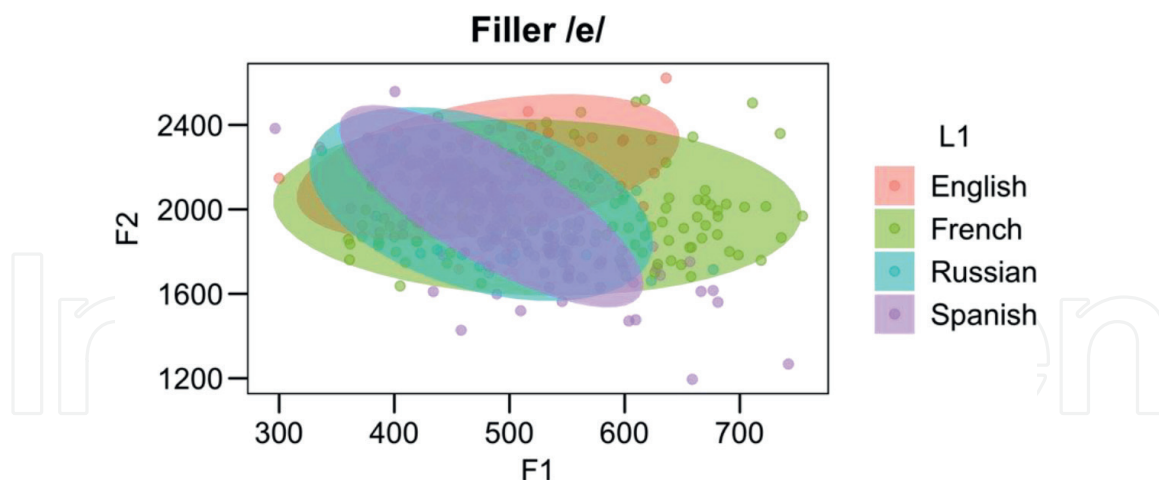


Figure 2.
Dispersion area of F1 and F2 values when the pause is filled by /e/.

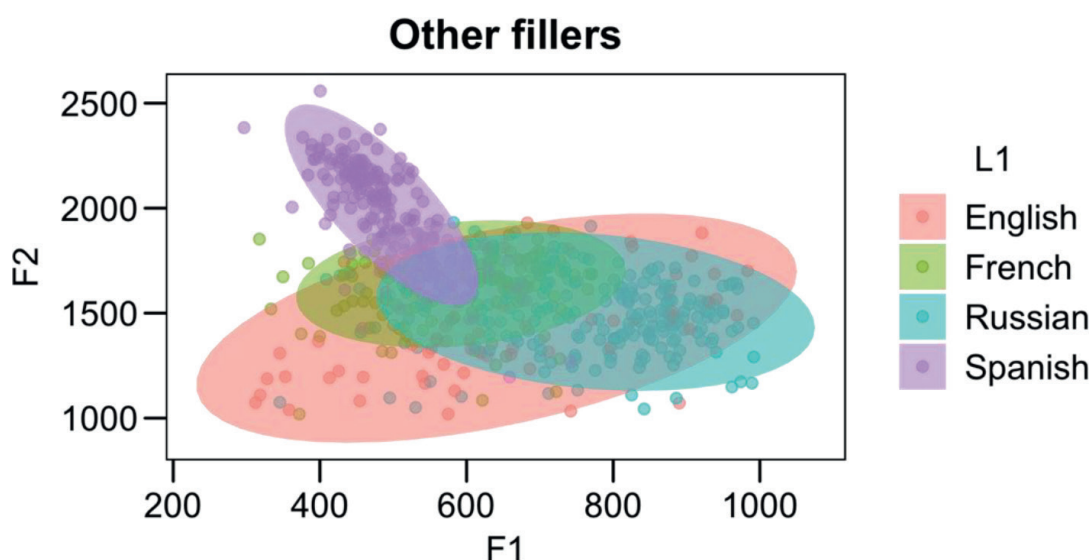


Figure 3.
Dispersion area of fillers categorized like non /e/ compared with native speakers (Spanish).

To address this question, a Chi-square test was conducted with two variables: proficiency and the phonetic category into which the filler was categorized. The test results revealed a significant correlation between these two variables. The higher the language proficiency, the more similar the filler is to the Spanish filler /e/. As **Figure 4** shows, 61.59% of advanced speakers tend to produce an /e/, while only 13.73% of intermediate speakers do so: $\chi^2_{\text{Pearson}}(1) = 163.36$, $p = 2.09e-37$, $V_{\text{Cramer}} = 0.50$, $CI_{95\%}[0.43, 1.00]$, $nobs = 659$.

4. Discussion and conclusions

The results obtained in this study indicate that foreign language learners tend to use a vocalic element similar to that of their native language. Examining the use of fillers in the context of a foreign language and determining the implications for teaching any foreign language are crucial for acquiring this new language and enhancing

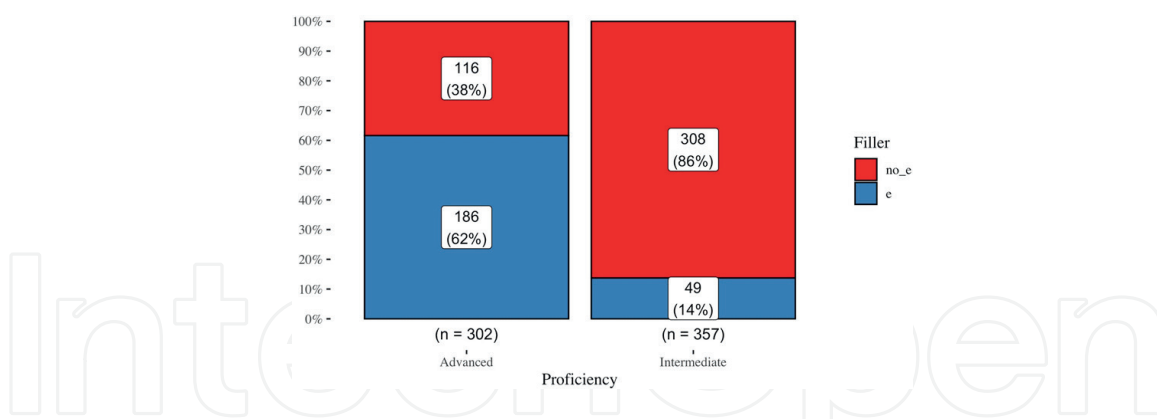


Figure 4.
Percentages of /e/ and non /e/ production considering proficiency of learners.

competence. Fitriati [43] concluded their paper about fillers with the following statement: “increasing EFL students’ awareness of the importance of fillers in verbal communication will help them improve their communication strategies as well as develop their communicative competence.” Using appropriate fillers when students are learning a new language increases their proficiency and diminishes the effect of a foreign accent. The first step is to vocalize fillers with the same vocalic element as native speakers. Although non-native speakers used both lexicalized and non-lexicalized filler in argumentative discourse [43], Navratilova [44] asserts that the most common filler used by male and female students in argumentative discourse was filled pauses unrelated to lexical units. Therefore, this type of filler is frequent enough to be considered in oral classes for foreign languages.

We have also observed differences between the two vocalic segments that they employ when they apply different communication strategies in the target language. When they produce a vowel sound similar to /e/, the acoustic characteristics are identical to native speakers. However, when they use the vowel from their native language, the significant differences in acoustic parameters between the two vocalizations for filled pauses demonstrate that they are completely different. Piccaluga [45] pondered whether the phonological system of vowels in various languages would affect the realization of filled pauses in trilingual speakers. Their finding does not suggest that speakers can alter the characteristics of their fillers when switching between languages. In foreign languages, more emphasis should be placed on teaching students how to use filled pauses; however, teachers must first recognize the significance of students knowing how to use them.

Furthermore, our data indicate that students with an advanced level produce a greater proportion of vocalic filled pauses equal to the language they are learning. F1 and F2 values highlight the phenomenon of language transfer from L1 to the new foreign language. In Ref. [39] also described this type of transfer in relation to Hungarian L1 learners of English. Intermediate and advanced proficiency are distinguished by the use of vocalic filled pauses. Most languages use schwa or central vowels to fill pauses [46], but the most common vocalic segment in Spanish is very close to /e/, implying that more articulatory effort is required to produce this segment. This method of filling in sound pauses distinguishes Spanish from other languages since it differs greatly.

As a conclusion and answering the questions formulated at the beginning, we can note that the acoustic parameters of filled pauses classified as /e/ are distinct from those categorized as not /e/. We assume that when students do not produce /e/, they fill the pause with a different Spanish vocalic element. Filled pauses categorized as

/e/ are similar to those used by native Spanish speakers and are independent of their L1 origin. The frequency region categorized as non-/e/ depends on the speaker's L1. French speakers use a vocalic segment placed in a central region, whereas Russian speakers use something more akin to an /a/. Due to the great variability, it is difficult to determine the frequency region in which English vocalic elements occur. The selection of these filled pauses results from the speaker's proficiency. Learners with a high proficiency were more likely to produce the correct filler than those with an intermediate level.

Finally, in view of our results, we can state what Blondet [5] had assumed, the best way to demonstrate that an individual has acquired a foreign language is for that person to produce pauses filled with the sound of the acquired language, especially if these pauses serve as identifiers of the language. Therefore, foreign language learners and instructors must be made aware of the significance of these filled pauses to enhance the learner's communicative competence [47].

IntechOpen


Author details

María J. Machuca

Autonomous University of Barcelona, Barcelona, Spain

*Address all correspondence to: mariajesus.machuca@uab.cat

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Candea M, Vasilescu I, Adda-Decker M. Inter- and intra-language acoustic analysis of autonomous fillers. In: Proceedings of 5th Disfluency in Spontaneous Speech Workshop (DiSS-05); 10-12 September 2005. Aix-en-Provence, France; 2005. pp. 47-51. Available from: https://www.ida.liu.se/~robek28/conferences/diss2017/DiSS2005_Proceedings.pdf
- [2] Cruttenden A. Intonation. Cambridge: Cambridge University Press; 1986. p. 214
- [3] Graham L. The case of *este* vs. *eh* in Latin American Spanish. *Spanish in Context*. 2018;15(1):1-26. DOI: 10.1075/sic.00001.gra
- [4] Rebollo L. Pausas y ritmo en la lengua oral. *Didáctica de la pronunciación*. In: Moreno F, Gil M, Alonso K, editors. *El español como lengua extranjera: del pasado al futuro*. Actas del VIII Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá; 1997. pp. 667-676
- [5] Blondet MA. Las pausas llenas: marcas de duda e identidad lingüística. *Lingua Americana*. 2001;8:5-15
- [6] Maclay H, Osgood CH. Hesitations phenomena in spontaneous English speech. *Word*. 1959;15:19-44. DOI: 10.1080/00437956.1959.11659682
- [7] Machuca MJ, Llisterri J, Ríos A. Las pausas sonoras y los alargamientos en español: un estudio preliminar. *Normas. Revista de Estudios Lingüísticos Hispánicos*. 2015;5:81-96
- [8] Nicholson H, Eberhard K, Scheutz M. “um...I don’t see any”: The function of filled pauses and repairs. In: Proceedings of 5th Workshop on Disfluency in Spontaneous Speech and 2nd International Symposium on Linguistic Patterns in Spontaneous Speech; 25-26 September 2010. Tokyo, Japan; 2010. pp. 89-92. Available from: https://www.isca-speech.org/archive/interspeech_2010/index.html
- [9] Wennerstrom A. Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics*. 1994;15(4):399-420
- [10] Cestero AM. El intercambio de turnos de habla en la conversación. *Análisis sociolingüístico*. Alcalá de Henares: Universidad de Alcalá de Henares, Servicio de Publicaciones; 2000 p. 308
- [11] Beňuš Š. Cognitive aspects of communicating information with conversational fillers in Slovak. In: Proceedings of 4th IEEE International Conference on Cognitive Infocommunications, Cog Info Com. December 2013. Budapest, Hungary; 2013. pp. 271-276. DOI: 10.1109/CogInfoCom.2013.6719255
- [12] Clark H, Fox TJ. Using *uh* and *um* in spontaneous speaking. *Cognition*. 2002;84:73-111
- [13] Stepanova S. Some features of filled hesitation pauses in spontaneous Russian. In: Trouvain J, Barry WJ, editors. *Proceedings of 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany. 2007. pp. 1325-1328
- [14] Goldman-Eisler F. *Psycholinguistics: Experiments in Spontaneous Speech*. Londres/Nueva York: Academic Press; 1968 169 p

- [15] Lickely RJ, Bard EG. On not recognizing disfluencies in dialogue. In: Proceedings of the Fourth International Conference on Spoken Language (ICSLP 96); October 1996. Vol. 3. Philadelphia; 1996. pp. 1876-1879. DOI: 10.1109/ICSLP.1996.607998. Available from: <http://www.asel.udel.edu/icslp/cdrom/about.htm>
- [16] Eriksson S. Localization, Frequency, and Functions of Filled Pauses: Five American Politicians' Use of "er" and "Erm" in the "Talk Show Larry King Live" [Thesis]. Finland: University of Turku; 2012
- [17] El TI. reconocimiento del habla. In: Llisterri J, Machuca MJ, editors. Los sistemas de diálogo. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45); 2006. pp. 81-98
- [18] Calò E, Rosemplat T. Detection of fillers in conversational speech. MSc Natural Language Processing M1 Supervised Project. Available from: https://idmc.univ-lorraine.fr/wp-content/uploads/2020/09/G3-Calo%CC%80_Rosemplat_fillers_realization_report.pdf. [Accessed: June 1, 2022]
- [19] Zhu G, Caceres JP, Salamon J. Filler Word Detection and Classification: A Dataset and Benchmark. 2022. Interspeech. Available from: <https://www.zhuanzhi.ai/paper/40272e60a2bd4cbac2d196b469ee6430>. [Accessed: June 1, 2022]
- [20] Lala D, Nakamura S, Kawahara T. Analysis of effect and timing of fillers in natural turn-taking. In: Proceedings of Interspeech 2019; September 2019. Graz, Austria; 2019. pp. 4175-4179. DOI: 10.21437/Interspeech.2019-1527. Available from: <https://dblp.org/rec/conf/interspeech/2019.html?view=bibtex>
- [21] Brander D. Phonetic characteristics of hesitation vowels in Swiss German and their use for forensic speaker identification. In: IAFPA 2014. 23rd Annual Conference of the International Association for Forensic Phonetics and Acoustics. Zurich, Switzerland; 2014. Available from: https://www.pholab.uzh.ch/static/IAFPA/IAFPA2014_Conference_Material.pdf
- [22] Machuca MJ, Llisterri J, Ríos A. Caracterización del hablante con fines judiciales: fenómenos fónicos propios del habla espontánea. In: Lingüística aplicada y transferencia del conocimiento: empleabilidad, internacionalización y retos sociales. Cádiz, Spain: 36º Congreso Internacional de la Asociación Española de Lingüística Aplicada; 19-21 April 2018; 2019
- [23] Braun A, Rosin A. On the speaker-specificity of hesitation markers. In: Proceedings of the 18th International Congress of Phonetic Sciences; 10-14 August 2015. London: International Phonetic Association; 2015. pp. 0731.1-0731.5. Available from: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/preface.html>
- [24] McDougall K, Duckworth M. Profiling fluency: An analysis of individual variation in disfluencies in adult males. *Speech Communication*. 2017;95:16-27
- [25] Künzel HJ. Some general phonetic and forensic aspects of speaking tempo. *The International Journal of Speech, Language and the Law*. 1997;4(1):48-83
- [26] Ishihara S, Kinoshita Y. Filler words as a speaker classification feature. In: Tabain M, Fletcher J, Grayden D,

- Hajek J, Butcher A, editors. Proceedings of Australasian International Conference on Speech Science and Technology. Melbourne, Australia. 2010. pp. 34-37. Available from: <https://assta.org/proceedings/sst/SST-10/WELCOME.PDF>
- [27] Cicres J. Comparación forense de voces mediante el análisis multidimensional de las pausas llenas. *Revista Signos. Estudios de Lingüística.* 2014;**47**(86):365-384
- [28] Council of Europe (CEFR). Common European Framework of Reference for Languages Learning, Teaching, Assessment. Cambridge: Cambridge University Press; 2001. p. 275
- [29] Giannini A. Hesitation phenomena in spontaneous Italian. In: Solé MJ, Recasens D, Romero J, editors. Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona: Adelaida: Casual Productions; 2003. pp. 2653-2656
- [30] Morin G, Tucker B. The acoustic characteristics of um and uh in spontaneous Canadian English. In: Proceedings of DiSS 2021; 25-27 August 2021. France: Paris 8 University; 2021
- [31] Hughes V, Wood S, Foulkes P. Strength of forensic voice comparison evidence from the acoustics of filled pauses. *International Journal of Speech Language and the Law.* 2016;**23**:99-132. DOI: 10.1558/ijssl.v23i1.29874
- [32] Fuchs S, Koenig L, Gerstenberg A. A longitudinal study of speech acoustics in older French females: Analysis of the filler particle euh across utterance positions. *Language.* 2021;**6**:211. DOI: 10.3390/languages6040211
- [33] Gil J, Lahoz-Bengoechea JM, Villa J. La vocal de relleno en español y en ruso: caracterización acústica e implicaciones teóricas. *Estudios Filológicos.* 2017;**60**:69-94. DOI: 10.4067/S0071-17132017000200004
- [34] Machuca MJ, Ríos A. Estructura formántica de las pausas sonoras en español. In: Fernández Planas AM, editor. 53 reflexiones sobre aspectos de la fonética y otros temas de lingüística. Barcelona: Laboratorio de Fonética de la Universidad de Barcelona; 2016. pp. 67-76
- [35] Villa J, Gil J, Lahoz-Bengoechea JM. Las vocales de relleno en español: nuevos datos y algunas reflexiones. In: Ruiz Mirayes L, Alvarez Silva MR, Muñoz Alvarado A, editors. Nuevos estudios sobre comunicación social. La Habana: Centro de Lingüística Aplicada; 2017. pp. 165-169
- [36] Rieger C. Idiosyncratic fillers in the speech of bilinguals. In: Proceedings of DiSS '01: Disfluency in Spontaneous Speech; 29-31 August 2001. Edinburgh; 2001. pp. 81-85. Available from: http://www.disfluency.org/app/download/29670881/DiSS2001_Proceedings.pdf
- [37] Khojasteh Rad S, Nadzimah Abdulla A. Effect of context on types of hesitation strategies used by Iranian EFL learners in L2 Oral language tests. *English Language Teaching.* 2012;**5**(7):102-109. DOI: 10.5539/elt.v5n7p102
- [38] Erten S. Teaching fillers and students' filler usage: A study conducted at ESOGU preparation school. *International Journal of Teaching and Education.* 2014;**2**(3):67-79
- [39] Gósy M, Gyarmathy D, Beke A. Phonetic analysis of filled pauses based on a Hungarian-English learner corpus. *International Journal of Learner Corpus*

Research. 2017;**3**:149-174. DOI: 10.1075/ijlcr.3.2.03gos

[40] Lahoz-Bengoechea JM, Gil J, García de León CL G. EMULANDO: Corpus de habla con acento no nativo auténtico y disimulado. In: Lahoz-Bengoechea JM, Pérez Ramón R, editors. Tools and Resources for Speech Sciences Málaga. Málaga, Spain: Universidad de Málaga; 2019. pp. 97-101

[41] Boersma P, Weenink D. Praat: doing phonetics by computer [Computer program]. Version 6.1.5.1. 2021. Available from: <http://www.praat.org/>

[42] Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015;**67**(1):1-48. DOI: 10.18637/jss.v067.i01

[43] Fitriati W, Mujiyanto J, Susilowati E, Melati P. The use of conversation fillers in English by Indonesian EFL Master's students. *Linguistic Research*. 2021;**38**(Special Edition):25-52. DOI: 10.17250/khisli.38..202109.002

[44] Navratilova L. Fillers used by male and female students of English education study program in argumentative talks. *Journal of English Linguistics and Language Teaching*. 2015;**2**(1):1-10

[45] Piccaluga M, Poch-olivé D, Harmegnies B. Effects of the multilingual phonologic competence on the phonetic properties of filled pauses. In: Sock R, Fuchs S, Laprie Y, editors. 8th International Seminar on Speech Production. Strasbourg, France. 2008. pp. 141-144. Available from: <https://www.leibniz-zas.de/en/research/publications/details/publications/3603-proceedings-of-the-8th-international>

[46] O'Shaughnessy D. Recognition of hesitations in spontaneous

speech. In: Proceedings of the 1992 IEEE International Conference on Acoustics, ICASSP'92, Speech and Signal Processing. San Francisco: IEEE Computer Society; 1992. pp. 521-524

[47] Basurto N, Hernández Alarcón MM, Mora I. Fillers and the development of Oral strategic competence in foreign language learning. *Porta Linguarum: Revista internacional de didáctica de las lenguas extranjeras*. 2016;**25**:191-201