FILLER PARTICLES IN ENGLISH AND SPANISH L1 AND L2 SPEECH

Beeke Muhlack

Language Science and Technology, Saarland University
muhlack@lst.uni-saarland.de

ABSTRACT

The filler particle (FP) distribution and vowel quality of 10 female English native speakers and 10 female native Spanish speakers were investigated both in their native language (L1) and their second language (L2). As expected, speakers produce more FPs in their L2 speech and speaker groups prefer different types of FPs. Native English speakers produce more vocalic-nasal FPs (*um*) than Spanish speakers while the latter produce more vocalic FPs (*uh*). The vowel quality in the FPs differs between the speakers' L1 (open central vowel for English and close-mid front vowel for Spanish) but vowel areas merge in the L2, which can be explained with the speakers' attempt to produce native-like FPs of the target language.

Keywords: filler particles, disfluencies, second language learning, English, Spanish

1. INTRODUCTION

The aim of this paper is to compare the production of filler particles (FPs), i.e., the frequency distribution and vowel quality, in English and Spanish as a native (L1) and second language (L2). Work on FPs and their phonetic characteristics has increased over the last decades. However, languages other than English (e.g., Spanish [1]) are still under-researched in this area. Relevant studies for Spanish FPs show the following: The preferred type of FP in Spanish seems to be the vocalic FP with approx. 76% of all FPs being realised as *uh*, 13% as nasal FP *hm* and only 11% with a vocalic-nasal FP *um* [2]. The investigation of FP usage and quality of 24 L1 Spanish speakers living in the US shows that the vowel quality of FPs in their L1 changes from [e] to a lower, more central variant ([ə] or [a]) with increasing use of English (proficiency and length of stay) [1].

Similarly, [2] showed that Spanish monolinguals use an FP vowel with a lower F1 and a higher F2 than Spanish-Afrikaans bilinguals. It seems that Spanish speakers produce a fronted and more close vowel in their FPs also compared to other languages [3]. In comparison with vowels in lexical material, [4] show that the FP vowel is closest to the realisations of /e/ but claim it is produced with a higher F1 and a lower F2 than the lexical vowels. Unfortunately, in [3, 4] the degree of overlap between the realisations of /e/ and the vowel in FPs was not reported.

From the numerous studies on FPs in English, for instance [5] reports the vocalic-nasal FP *um* as the preferred type for British English speakers (81% *um*, 18% *uh*). The vowel quality of English FPs is often reported to vary between the central vowels [ə] and [ʌ][1] and the back vowel [ɑ] [6]. The vowels in the FPs usually spread over a large part of the central vowel space [7]. Results in [8] suggest that the vowel in FPs in English is closest (in terms of smallest Euclidean distance) to the open-mid central vowel [ʌ], although the English speaker group is quite heterogeneous in terms of their English accents (which may introduce further variation).

Work on L2 learning has shown that the use of disfluencies, and thus also FPs, is typically higher in an L2 [9, 10, 11, 12] but decreases for pauses with rising L2 proficiency [13]. An L2 effect for pauses but not for FPs was also found by [14] for advanced English learners of German. The preference of the vocalic FP in L2 English (intermediate and advanced learners) by native Spanish speakers is reported by [15]. This seems to suggest that learners transfer their native FPs to their L2. The results in [16] show that advanced learners of an L2 (English and German) are able to adapt the vowel quality of the target language in their FPs but intermediate learners are not. The latter group seems to transfer their L1 FPs to the L2.

Furthermore, [17] found that German-French bilinguals produce distinct vowel qualities in the two languages but that the weaker language would show a shift towards the vowel quality of the dominant language. For Afrikaans-Spanish bilinguals from Patagonia, Argentina, [2] found that the speakers do not share the vowel quality of Afrikaans and Spanish monolinguals, but they still produce separate vowel qualities for their FPs in both languages.

Based on the reviewed literature, it is assumed that English native speakers prefer the vocalic-nasal FP *um* while Spanish native speakers prefer the vocalic FP *uh*. It is expected that language

learners also employ their preferred FPs in their L2 [15]. Distinct vowel qualities for each language are expected to be observed and, furthermore, a shift towards the L2 vowel quality is expected in the second language context.

2. MATERIAL

For this study, a subset of the Diapix-FL corpus [18] is used, which consists of 20 female speakers (10 native English, 10 native Spanish/Basque) completing a dialogue task in both their L1 and L2 (Spanish and English, respectively). Native speakers of English were recorded at the University of Edinburgh, native Spanish/Basque speakers were recorded at the University of the Basque Country in Vitoria. The speakers are quite proficient in their L2 (B2-C1 on CEFR[2]) as both groups study their L2 in their second year at university. The native Spanish participants all passed a B2 proficiency test for English as part of their study programme to reach their second year, native English speakers self-reported their proficiency level for Spanish as their L2. Four male speakers (2 English, 2 Spanish) from the corpus were excluded in order to keep the speaker group as homogeneous as possible.

The speakers were grouped in same-L1 pairs and given a spot-the-difference task in both L1 and L2. Each speaker was given a picture, different versions of a scene, for which they had to cooperate to find 12 differences without being able to see each other's pictures. Dialogue partners solved three spot-the-difference tasks per language. Each speaker was recorded on a separate channel. For more details on the recording method, see [19, 20].[3]

Orthographic annotations are included in the corpus along with annotations of silent pauses (including breath noises), elongations, and FPs. All FPs are marked with one symbol, so for the re-annotation distinct labels for vocalic (*uh*), nasal (*hm*) and vocalic-nasal (*um*) FPs were used. Additionally, ten tokens (annotated in stressed positions) of each corner vowel (/i u a/ for Spanish and /i u ɑ/ for English) for all L1 speech were annotated along with ten tokens of one additional vowel that is frequently reported to occur in the FPs of the respective language (/e/ for Spanish, /ʌ/ for English).

3. RESULTS

A total of 2,737 FPs was found in the subset of the corpus: 245 nasal FPs (*hm*), 1,118 vocalic FPs (*uh*) and 1,374 vocalic-nasal FPs (*um*).

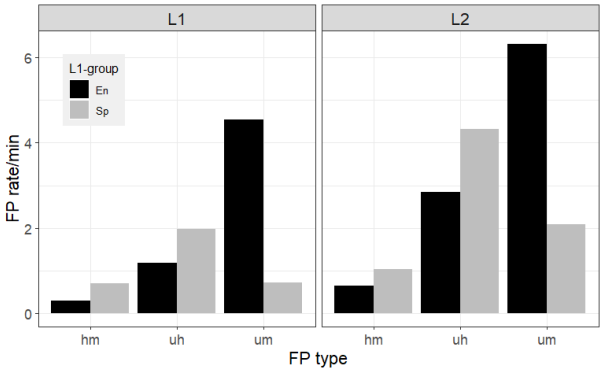

Figure 1: FP rate per minute in L1 and L2 speech by both speaker groups. Native English speakers in black, native Spanish speakers in grey.

3.1. Frequency distribution

FPs are more frequent in each L2 than in the respective L1 speech. L1 English shows an overall FP rate of 1.9 per minute while L1 Spanish shows a lower rate with 1.0 FP per minute. Native English speakers increase their FP rate only slightly in their L2 (2.2 FPs/min for L2 Spanish), whereas native Spanish speakers show a rate of 2.5 FPs per minute in their L2, which is considerably higher compared to their low L1 rate. The frequency distribution of the four FP types in both speaker groups (English and Spanish native speakers) and both conditions (L1, L2) is shown in Figure 1. In their respective native languages (left panel), L1 Spanish speakers use more vocalic FPs than the other FP types, while L1 English speakers use the vocalic-nasal FP type *um* more frequently (4.8 vs. 0.7 *um*/min). The nasal FP *hm* is used rarely but more often by Spanish speakers. The same pattern is also visible in L2 speech (right panel), with numbers increasing overall.

During the classification of the FPs, an observation was made: speakers, when producing a vocalic-nasal FP, do not always use the bilabial nasal /m/ but also an alveolar /n/ in some instances. An auditory inspection by one annotator (the author) revealed that 3.4% (i.e., 38/1,108) of all vocalic-nasal FPs of the English native speakers (in both L1 and L2) were produced with a nasal that was not bilabial and 18.1% (i.e., 48/266) of those by the Spanish native speakers were produced with a nasal that was not bilabial. The majority of non-bilabial nasals were alveolar, while a few were also categorised as labiodental, velar or a sequence of alveolar and bilabial nasals. To shed more light on the phenomenon, the annotation of several annotators should be taken into account.

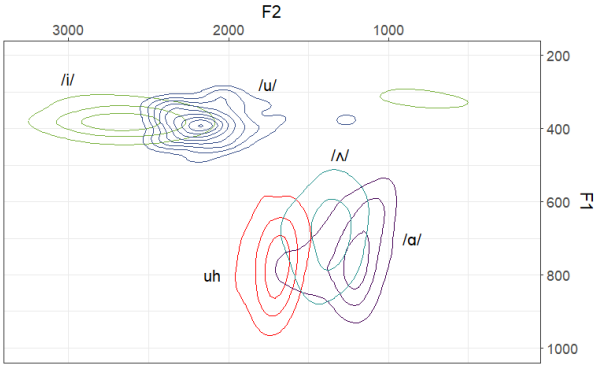

Figure 2: Formant values of realisations of *uh* (red) in English L1 compared to those of the lexical vowels /ɑ i u ʌ/.

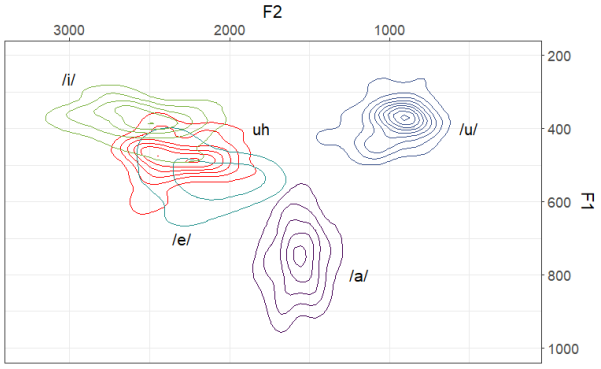

Figure 3: Formant values of realisations of *uh* (red) in Spanish L1 compared to those of the lexical vowels /a i e u/.

3.2. Vowel quality

The first and second formant of all purely vocalic FPs and the corner vowels of L1 English and L1 Spanish were measured in Praat [21] at the midpoint of the vowel using the Burg method.[4] To visualise the vowel quality produced in the FP *uh*, two-dimensional (F1, F2) kernel density distributions of the lexical vowels and the FP vowel were plotted in R [22] using ggplot2 [23].

Figure 2 shows that the vowel in the native English FPs partly overlaps with the realisations of the lexical vowels (also L1 English) /ɑ/ and /ʌ/. Based on this observation, the overlap in the two-dimensional vowel spaces between these two lexical vowels and the FP vowel was calculated using the Pillai-score [24], that ranges from 1 (no overlap) to 0 (complete overlap). The vowel quality of the FP *uh* overlaps more with the open-mid vowel /ʌ/ (Pillai: 0.27) than the open vowel /ɑ/ (Pillai: 0.46). The /u/-fronting has been described before for the same data set [19].

Figure 3 shows that the vowel in the native Spanish FPs overlaps with the two front vowels /i/ and /e/ of native Spanish speech to a considerable degree. The Pillai scores show more overlap with the close-mid vowel /e/ (Pillai: 0.22) than the close vowel /i/ (Pillai: 0.34).

In order to answer the question whether speakers transfer the vowel quality of their L1 FP to their L2, Figure 4 compares the FP *uh* by both speaker groups in their L1 and L2 speech. The left panel shows that FPs in English and Spanish L1 are produced with distinct vowel qualities (Pillai: 0.67). The L2 FPs (right panel) show vowel qualities that spread over a larger area of the central vowel space for both speaker groups. However, they also show a bimodal distribution for both groups. Some FPs in L2 Spanish by the English speakers are produced with a lower F1 and a higher F2 while the majority of FPs is produced with the same vowel quality as

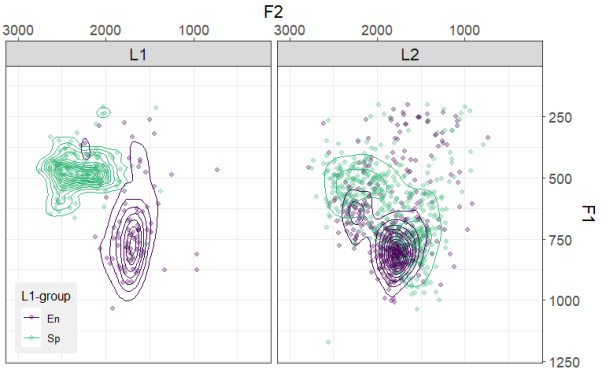

Figure 4: Vowel quality of *uh* in L1 and L2 speech by native English speakers (purple) and native Spanish speakers (green).

observed in their L1. Conversely, for L2 English by the Spanish speakers some of the FPs are produced with a native-like vowel quality (similar to the close-mid /e/) but a large portion is produced with a higher F1 and a lower F2, approximating the native English vowel quality.

4. DISCUSSION

The results of the current study confirm that L1 English prefers the vocalic-nasal FP (*um*) and L1 Spanish prefers the vocalic FP (*uh*) [2, 5]. Nasal FPs are not very frequent in either speaker group but for L1 Spanish *hm* is as infrequent as *um*. The frequency of FPs in L2 speech increases as reported before [9, 10, 11, 12], and moreover, the preferred type of FP remains constant [15] with native Spanish speakers preferring the vocalic FP *uh* and the native English speakers preferring the vocalic-nasal type *um* in both their L1 and L2. A reason for the preference of the specific FP type may be the syllable structures in both languages. In English, closed syllables seem to be more common than open syllables as observed from the data reported in [25], while in Spanish closed syllables do occur, but most often word-internally [26]. Another reason for the

preference of *uh* over *um* may be the restriction for Spanish word-final closed syllables which only allows the consonants /l, n, r, s, d/ in the coda[5] [26]. This restriction would also explain the occurrence of alveolar nasals in the vocalic-nasal FPs of the native Spanish speakers, yet not those produced by the native English speakers.

The results of the comparison of vowel qualities between lexical vowels and the FP *uh* in English show that the realisations of the open-mid central vowel /ʌ/ best represent the vowel of the FP, although vowel spaces do not entirely overlap. The FP vowel in English is produced more centrally than the vowel /ʌ/. The Spanish FP vowel, on the other hand, is produced as a close-mid front vowel, similar to the realisations of the Spanish lexical vowel /e/. In Spanish, [e] and [ɛ] are allophones of the phoneme /e/ [26], so the annotated corner vowels in Spanish include tokens of both allophones. The FP vowel overlaps with (the upper) half of the vowel space covered by the vowel phoneme /e/, suggesting that the vowel quality in *uh* is represented by [e] rather than [ɛ]. This is also supported by previous literature [3, 2]. However, the claim by [4] that the Spanish FP is produced with a lower F2 is not supported by the data presented here. Although a higher F1 can be observed, this may be due to the inclusion of the allophone [ɛ] in the set of corner vowels. Note that native Spanish speakers were bilinguals in Spanish and Basque, which may have influenced the vowel quality of the FPs. However, as the vowel systems of Spanish and Basque are the same [27, 26] this influence is assumed to be rather small.

The comparison between L1 and L2 speech shows that while the FP vowels in L1 English and Spanish are very distinct, this cannot be observed in the speakers' L2 speech. The distributions of FP vowels of Spanish and English L1 speakers, producing FPs in their L2 show two peaks, which suggests that some hesitations approximate the realisations of Spanish /e/ and some approximate the realisations of English /ʌ/. It seems that advanced learners are able to adapt the FP's vowel quality, but this is not the case for all the data. The results suggest that L1 Spanish speakers may be better at approximating the vowel quality of English FPs and that L1 English speakers are more likely to keep their native FP in their L2. It is likely that individual speakers adapt the vowel quality of the target language while others keep their native vowel quality. Factors that may influence the switch from L1 to L2 FP vowel quality may be L2 proficiency, the exposure to the target language by native speakers, or the speakers' stance towards the foreign language. Another possibility is that speakers do not switch from one vowel quality to the other entirely, but they employ a mixture of native-like and foreign FPs. Further investigations into speaker-specific patterns are planned for future work.

5. CONCLUSION

To conclude, this study has shown that filler particles (FPs) in English and Spanish native speech are quite distinct in (i) the type that is preferred (*um* vs. *uh*) and (ii) the vowel quality that is used approximating the realisations of the lexical vowels /ʌ/ in English and /e/ in Spanish. Furthermore, it is shown that learners of an L2 are able to produce FPs with a native-like vowel quality, even though a full adaptation was not observed. It is assumed that the degree of target language FP adaptation is speaker-specific, but this hypothesis is in need of further investigation.

The results discussed here offer insights into the realisations of FPs in two languages, suggesting that different languages employ specific FP paradigms. The tendency that native-like qualities of FPs (their vowel quality and preferred type) transfer to the speaker's L2 even for advanced learners may be highly relevant in fields where the speaker's background is unknown (forensic phonetic casework, language analysis for determination of origin). The results also support the view that disfluencies should be discussed in the L2 classroom to raise the learners' awareness of the foreign FP realisations. Whether the discussion of FPs in the L2 context helps the learners better approximate the L2 FPs is still an open question in L2 research.

[1] In phonetic transcription for English the symbol ʌ is commonly used to describe an open-mid central vowel [28] as opposed to a back vowel as described by the IPA.
[2] The Common European Framework of Reference for Languages (CEFR) [29] provides guidelines for the categorisation of L2-fluency and proficiency.
[3] The DIAPIX-FL corpus is freely available at http://datashare.is.ed.ac.uk/handle/10283/346
[4] Maximum formant: 5.5 kHz, max. number of formants: 5, window length: 0.025 s, dynamic range: 50 Hz
[5] Other consonants may occur in loanwords.

6. REFERENCES

[1] D. Erker and J. Bruso, "Uh, bueno, em...: Filled pauses as a site of contact-induced change in Boston Spanish," *Language Variation and Change*, vol. 29, no. 2, pp. 205–244, 2017.

[2] L. Garciá-Amaya and S. Lang, "Filled pauses are susceptible to cross-language phonetic influence,"

Studies in Second Language Acquisition, pp. 1–29, 2020.

[3] M. Candea, I. Vasilescu, and M. Adda-Decker, "Inter- and intra-language acoustic analysis of autonomous fillers," in *Workshop on Disfluency in Spontaneous Speech Workshop (DiSS 2008), Aix-en-Provence*, 2008, pp. 47–52.

[4] I. Vasilescu, R. Nemoto, and M. Adda-Decker, "Vocalic hesitations vs vocalic systems : a cross-language comparison," *Proceedings of the 16th International Congress of Phonetic Sciences*, no. August, pp. 1101–1104, 2007.

[5] E. de Leeuw, "Hesitation markers in English, German, and Dutch," *Journal of Germanic Linguistics*, vol. 19, no. 2, pp. 85–114, 2007.

[6] E. Shriberg, "Preliminaries to a theory of speech disfluencies," Ph.D. dissertation, 1994. [Online]. Available: ftp://130.107.33.205/pub/papers/shriberg-thesis.pdf

[7] V. Hughes, S. Wood, and P. Foulkes, "Strength of forensic voice comparison evidence from the acoustics of filled pauses," *International Journal of Speech, Language and the Law*, vol. 23, no. 1, pp. 99–132, 2016.

[8] B. Muhlack, "The vowel quality of non-lexical hesitation particles in German and English L1 and L2 speech [Poster]," in *Phonetik und Phonologie im deutschsprachigen Raum (P&P16), Trier*, 2020.

[9] C. Brand and S. Götz, "Fluency versus accuracy in advanced spoken learner language," *International Journal of Corpus Linguistics*, vol. 16, no. 2, pp. 255–275, 2011.

[10] G. Gilquin, "Hesitation markers among EFL learners: Pragmatic deficiency or difference?" in *Pragmatics and Corpus Linguistics: A Mutualistic Entente*, J. Romero-Trillo, Ed. Berlin, Heidelberg, New York: Mouton de Gruyter, 2008, pp. 119–149.

[11] L. Temple, "Second language learner speech production," *Studia Linguistica*, vol. 54, no. 2, pp. 288–297, 2000.

[12] R. Wiese, "Language production in foreign and native languages: Same or different?" in *Second Language Productions*, H. W. Dechert, D. Möhle, and M. Raupach, Eds. Tübingen: Gunter Narr Verlag, 1984, pp. 11–25.

[13] A. Riazantseva, "Second lanuage proficiency and pausing: A study of Russian speakers of English," *Studies in Second Language Acquisition*, vol. 23, no. 4, pp. 497–526, 2001.

[14] M. Belz, S. Sauer, A. Lüdeling, and C. Mooshammer, "Fluently disfluent?: Pauses and repairs of advanced learners and native speakers of German," *International Journal of Learner Corpus Research*, vol. 3, no. 2, pp. 118–148, 2017.

[15] J. Cenoz, "Pauses and hesitation phenomena in second language production," *ITL - International Journal of Applied Linguistics*, vol. 127-128, no. January 2000, pp. 53–69, 2000.

[16] B. Muhlack, "L1 and L2 production of non-lexical hesitation particles of German and English native speakers," in *Workshop on Laughter and Other Non-Verbal Vocalisations, Bielefeld*, 2020, pp. 44–47.

[17] J. J. Lo, "Between äh(m) and euh(m): The distribution and realization of filled pauses in the speech of German-French simultaneous bilinguals," *Language and Speech*, vol. 63, no. 4, pp. 746–768, 2020.

[18] M. Cooke, M. L. Garcia Lecumberri, and M. Wester, *DiapixFL*. LISTA Consortium: (i) Language and Speech Lab, Universidad del Pais Vasco, Spain and Ikerbasque, Spain; (ii) CSTR, University of Edinburgh, UK; (iii) KTH Royal Institute of Technology, Sweden; (iv) Institute of Computer Science, FORT, 2013.

[19] M. L. Garcia Lecumberri, M. Cooke, and M. Wester, "A bi-directional task-based corpus of learners' conversational speech," *International Journal of Learner Corpus Research*, vol. 3, no. 2, pp. 175–195, 2017.

[20] M. Wester, M. L. García Lecumberri, and M. Cooke, "DIAPIX-FL: A symmetric corpus of problem-solving dialogues in first and second languages," in *Interspeech, Singapore*, 2014, pp. 509–513.

[21] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," 2022. [Online]. Available: http://www.praat.org

[22] R Core Team, "R: A Language and Environment for Statistical Computing," in *R Foundation for Statistical Computing*. Vienna, Austria: R version 4.1.3, 2022. [Online]. Available: https://www.r-project.org/

[23] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. [Online]. Available: https://ggplot2.tidyverse.org

[24] M. C. Kelley and B. V. Tucker, "A comparison of four vowel overlap measures," *The Journal of the Acoustical Society of America*, vol. 147, no. 1, pp. 137–145, 2020.

[25] T. H. Crystal and A. S. House, "Articulation rate and the duration of syllables and stress groups in connected speech," *Journal of the Acoustical Society of America*, vol. 88, no. 1, pp. 101–112, 1990.

[26] C. Gabriel, "Phonetik und Phonologie des Spanischen," in *Linguistik im Sprachvergleich: Germanistik, Romanistik, Anglistik*, R. Klabunde, W. Mihatsch, and S. Dipper, Eds. Berlin: J.B. Metzler, 2022, ch. 2, pp. 27–48.

[27] J. I. Hualde and J. Ortiz de Urbina, *A Grammar of Basque*. Berlin, New York: Mouton de Gruyter, 2003.

[28] P. J. Roach, *English Phonetics and Phonology: A practical course*, 4th ed. Cambridge, New York, Melbourne: Cambridge University Press, 2009.

[29] Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Strasbourg: Cambridge University Press, 2011. [Online]. Available: https://rm.coe.int/1680459f97