

The following paper was presented at The 9th Workshop on Disfluency in Spontaneous Speech (DiSS 2019) held at ELTE Eötvös Loránd University in Budapest, Hungary on 12–13 September, 2019.

Title: Hesitation patterns in the Spanish spontaneous speech of Hungarian learners of Spanish

Author(s): Kata Baditzné Pálvölgyi

Abstract: This paper examines what native Spanish speakers find most disturbing in the pronunciation of Hungarian language learners of Spanish. Former research (Baditzné Pálvölgyi, 2019) showed that in spontaneous Spanish speech of at least threshold level Hungarian learners, one of the aspects that Spanish native speakers least tolerated was the way Hungarians hesitated. So the present paper focuses primarily on hesitation phenomena—lengthening and filled pauses—assuming that Hungarians hesitate more, and the lengthened segments are longer than the Spanish ones. In order to validate the hypothesis, an investigation comparing a corpus of Northern Spanish spontaneous speech to another corpus of advanced Hungarian learners of Spanish was conducted.

DOI: <https://doi.org/10.21862/diss-09-010-badi>

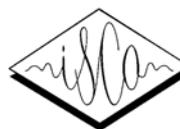
Citation (JIPA): Baditzné Pálvölgyi, Kata. 2019. Hesitation patterns in the Spanish spontaneous speech of Hungarian learners of Spanish. In: R. L. Rose & R. Eklund (eds.), *Proceedings of DiSS 2019, The 9th Workshop on Disfluency in Spontaneous Speech*, 12–13 September, 2019, Budapest, Hungary, 35–38.

The complete proceedings for DiSS 2019 are available as follows.

ISBN: 978-963-489-063-8

DOI: <https://doi.org/10.21862/diss-09>

DiSS 2019 was sponsored by The Faculty of Humanities, ELTE Eötvös Loránd University and the International Speech Communication Association (ISCA).



Hesitation patterns in the Spanish spontaneous speech of Hungarian learners of Spanish

Kata Baditzné Pálvölgyi

Department of Spanish Language and Literature, Eötvös Loránd University, Budapest, Hungary

Abstract

This paper examines what native Spanish speakers find most disturbing in the pronunciation of Hungarian language learners of Spanish. Former research (Baditzné Pálvölgyi, 2019) showed that in spontaneous Spanish speech of at least threshold level Hungarian learners, one of the aspects that Spanish native speakers least tolerated was the way Hungarians hesitated. So the present paper focuses primarily on hesitation phenomena—lengthening and filled pauses—assuming that Hungarians hesitate more, and the lengthened segments are longer than the Spanish ones. In order to validate the hypothesis, an investigation comparing a corpus of Northern Spanish spontaneous speech to another corpus of advanced Hungarian learners of Spanish was conducted.

Introduction

The topic of L1 disfluency has been widely discussed in numerous papers, but L2 disfluency phenomena should deserve more attention in research projects (Rieger, 2003) and also in the foreign language classroom (Belz et al., 2017).

According to Medgyes (2001), there are certain areas which, even if language learners stay in the target language country for a long time, they cannot master perfectly. These areas include vocabulary, idiomatic expressions, listening skills, fluency and pronunciation. If we concentrate on these latter two, we find that hesitation is connected to both fields. The sounds people employ when hesitating are not necessarily universal; and if hesitation is defined as the interruption of continuous speech, then it can be seen as a blocker of fluency. On the other hand, hesitation can also guarantee that the speaker will hold the dialogue turn, or at least attract listeners' attention (Bosker et al., 2015), so its role in spoken discourse is twofold. The question of holding turns is of considerable importance in Spanish, as this language is characterized by apparently violent rules concerning debate techniques, which may even cause culture shock to, for example, Hungarians, who are not used to the so-called Mediterranean debate. The ways of holding, and especially obtaining conversational turns seem too vehement for an outsider, and if learners of Spanish are not equipped

with the right strategies, they might lose ground in spontaneous discussions with native Spanish speakers, which is obviously not their goal.

Thus we must pay special attention to give learners of Spanish the necessary devices in order to make them self-confident participants in Spanish conversations and debates. As has been mentioned, hesitation schemes can be the clue to assure that the speaker can hold a conversational turn. That is why it is really important that language learners acquire the right hesitation patterns when learning the target language. But do Hungarians learn how to hesitate in Spanish? Even at advanced level, apparently not; earlier research (Baditzné Pálvölgyi, 2019) has shown that Spanish native speakers least tolerated the following aspects in the Spanish spontaneous speech of minimum B1 level Hungarian language learners: the pronunciation of certain vowels and consonants (mostly sibilants); the uncommon intonational patterns employed; the slowness of speech rate, as well as the way Hungarian learners hesitated.

If we examine the differences between Hungarian and Spanish as far as hesitation techniques are concerned, we find that there is a great difference: while Hungarians mostly use the schwa and the [m] to hesitate (Horváth, 2014), the Spanish apply the Spanish vowel [e], as well as [a] and the consonant [m] (Garrido Almiñana et al., 2017). This implies, also based on what native Spanish speakers said about the Spanish pronunciation of Hungarians—i.e. that their speech rate is too slow and they don't hesitate properly—that Hungarians hesitate differently because of mainly two reasons: first, the way they hesitate (tending to use [ə] instead of [e], for instance), and second, the quantity of their hesitation phenomena in their speech as compared to Spanish. This paper focuses on the second problem, formulating the following hypotheses: when hesitating, Hungarian learners of Spanish will:

- (1) tend to lengthen segments longer;
- (2) tend to hesitate more

than native Spanish people do.

So as to confirm these hypotheses, I carried out a two-step research, first analysing hesitation patterns in a native Spanish corpus, and then contrasting the results with phenomena attested in a corpus provided

by Hungarian learners of Spanish. The following sections will present the most important findings of the investigation.

The corpus

In order to validate the hypotheses, data was collected from both native Spanish speakers and Hungarian learners of Spanish. The corpus used in this study consisted of audio recordings coming from 25 Northern Spanish speakers (from Asturias, aged 18–23, all learning Humanities at the University of Oviedo), and from 25 Hungarian learners of Spanish (from Budapest, aged 19–25, all learning Spanish Philology at the Eötvös Loránd University). As for the Spanish competence of Hungarian learners, they all reached at least level B1 (‘threshold’) according to the CEFRL, and had spent no longer than a period of 4 months in Spain (students who had lived in Latin America were excluded from the corpus).

The recordings were conducted by Zoltán Kristóf Gaál (ELTE University), who interviewed the informants within the framework of a Map Task activity (the interviewees had to inform the interviewer about the correct itinerary according to a map). This guaranteed that the recorded speech was spontaneous, and hence suitable for an analysis from the perspective of hesitation phenomena. As the Hungarian informants were not native speakers of Spanish, it took them more time to express themselves and to explain the route. This is why, the total duration of the recordings in case of the language learners was almost double the duration of the native Spanish corpus.

Method

Hesitation is considered to be a disfluency phenomenon. There exist several subcategorizations connected to disfluency phenomena (Neuberger, 2014: 23), but Gósy’s (2002) was chosen. She differentiates disfluencies due to erroneous realization from disfluencies due to speaker insecurity. Phenomena related to this latter group include silent pauses, filled pauses (these are properly referred to as hesitations), fillers, repetitions, false starts and lengthenings. As both filled pauses and lengthenings may be considered ‘hesitations’ in a non-academic terminology, this research focuses on these two disfluency phenomena.

The following aspects were examined in the corpus:

(1) the frequency and duration of lengthened syllables (in case of lexemes), and of filled pauses (in case of vocalized hesitations); and

(2) the proportion of the time dedicated to hesitation phenomena as compared to the total speech time.

Annotation was carried out based on information extracted from the acoustic analysis software Praat (Boersma & Weenink, 2019), bearing in mind the following principle: when measuring lengthening, the syllable was always taken as the basic domain of analysis. This was so in case of the typical Spanish phenomenon of resyllabification (i.e. when a word-final consonant is resyllabified as the onset of the next word starting with a vowel, such as in [la: alondras], ‘the larks’), syllable length was only measured for [la:], as the *s* was resyllabified to the subsequent syllable *a-* in *alondras*.

Results

In order to validate the hypotheses, I examined (a) the percentage of hesitation/lengthening time as compared to the whole speech; (b) the frequency of hesitation/lengthening phenomena per minute and (c) the average duration of hesitations and lengthenings in the case of the 50 informants. Table 1. sums up the results.

Table 1. The characteristics of lengthenings and hesitations in both corpora.

Characteristics of lengthenings and hesitations	Hungarian	Spanish
hesitation and lengthening (%) of the whole speech, mean of 25 speakers	16.93	10.37
Std. dev.	7.64	6.22
Results of <i>t</i> -test for Difference of Means (equal variances)	Sig. (1-tailed) 0.00084; $p < 0.05$	
number of lengthening and hesitation phenomena (total)	2092	478
number of lengthening and hesitation phenomena per minute, mean of 25 speakers	20.21	11.22
Std. dev.	6.83	6.51
Results of <i>t</i> -test for Difference of Means (equal variances)	Sig. (1-tailed) 0.000009; $p < 0.05$	
duration of hesitation and lengthening phenomena (mean of 25 speakers)	0.51s	0.55s
Std. dev.	0.73	0.82
Results of <i>t</i> -test for Equality of Means (equal variances)	Sig. (2-tailed) 0.058; $p > 0.05$	

According to the first hypothesis, Hungarians tend to hesitate more than Spanish do, so I examined the proportion of hesitation phenomena first, that is, the time of lengthening and hesitation produced by

Hungarians compared to their total speech time. Hungarians effectively hesitated in a higher proportion of their total speech time than the Spanish informants did, hesitating during the 16.93% of their whole speech time as opposed to the 10.37% in the Spanish informants' case. These are mean values in case of the 25–25 informants, and statistic testing (one-tailed *t*-test for the difference of means) has revealed that there is a statistical difference between the means at the 95% confidence interval ($p < 0.05$).

This indicates that Hungarians hesitate in a higher proportion of their speech than Spanish do, and this can be due to two factors: the frequency and the duration of their hesitation phenomena. These two aspects were analysed, and as the data show, the difference between Hungarians and Spanish does not lie in the duration but in the frequency of hesitation. Almost four times more hesitation and lengthening phenomena was found in the Hungarian corpus than in the Spanish one. This number itself is impressive but we should bear in mind that Hungarians spoke longer, so the hesitation and lengthening phenomena per minute were analysed in both corpora. The mean values in both corpora were different, so a one-tailed *t*-test was applied again which proved that the difference of the two means—20.21 phenomena per minute in the case of Hungarians as opposed to 11.22 in the Spanish corpus—is statistically different at the 95% confidence interval ($p < 0.05$).

Average duration of hesitation phenomena, on the other hand, seemed no different in the two groups: surprisingly, Hungarians did not produce longer lengthenings compared to native Spanish speakers. The average duration of a hesitation or a lengthening in the case of the Spanish speakers was 0.55 s and 0.51 s in case of the Hungarian speakers, which difference, according to the two-tailed *t*-test for equality of means, is statistically not significant at the 95% confidence interval ($p > 0.05$).

The data confirm the first hypothesis: threshold level Hungarian speakers of Spanish tend to hesitate more than native Spanish speakers, which is due the higher frequency of hesitation phenomena but not to their duration. This also implies that the second hypothesis was not confirmed: segments affected by hesitation phenomena are not considerably longer in case of at least B1 CEFRL level Hungarian learners of Spanish as compared to native Spanish realizations.

Conclusion and discussion

As a reflection on the hypotheses, the data show that, as compared to native Spanish patterns, Hungarian learners of Spanish

(1) do not produce considerably longer lengthened segments than Spanish, but

(2) do hesitate more often.

A further aspect worth considering is a comparison between Hungarian students learning Spanish in foreign language classrooms in a non-immersion context and students learning Spanish abroad. According to the results of [García-Amaya \(2015\)](#) in case of learning Spanish as a foreign language, immersion study context yielded greater fluency and language proficiency improvement but also an increase in the use of filled pauses, for instance.

This raises the question whether the intolerance experimented on behalf of Spanish native speakers when judging Hungarians' Spanish—i.e. that they criticized Hungarian learners' hesitation patterns—was due to the relatively high frequency of hesitation phenomena or rather to other factors, such as the quality (like uncommon vocalic realizations resulting from negative transfer from Hungarian) or the position of hesitation phenomena.

Naturally, the present research is based on data interpretation coming only from Northern Spanish, so in the future further studies are encouraged to analyse hesitation phenomena in other dialects of Spanish as well.

Concerning the duration patterns of speaker-independent hesitation, in the future, it is convenient to stick to an objective model to see how lengthening is realized as compared to its context. [Cantero Serena's \(2019\)](#) duration standardization model would make it possible to see to what extent one lengthened syllable is longer than other syllabic intervals within the same utterance, for instance. Handling relative duration could help us generalize better than if we merely compare absolute duration data.

After defining areas to develop concerning hesitation phenomena, a further question is how to help Hungarian students to improve their hesitation techniques in Spanish in an effective way when no immersion language learning is feasible.

Acknowledgements

Supported by the ÚNKP-19-4 New National Excellence Program of the National Research, Development and Innovation Office. Sincere gratitude also goes to Zoltán Kristóf Gaál, Ph.D. student of the Romance Linguistics Doctoral Programme at Eötvös Loránd University, Budapest, for providing the recordings that were used for setting up the corpora. Also, special thanks to László Pálvölgyi and Péter Pálvölgyi, for their valuable comments on data interpretation.

References

- Baditzné Pálvölgyi, K. 2019. Magyar ajkú spanyol nyelvtanulók kiejtése spanyol anyanyelvűek szemével [The pronunciation of Hungarian-speaking Spanish learners through the eyes of native Spanish speakers]. In: *Magyar Alkalmazott Nyelvészeti Kongresszus (MANYE) XXVII*, 15–16 April 2019, Károli Gáspár Reformed University, Budapest, Hungary, 44–44.
- Belz, M., S. Sauer, A. Lüdeling & C. Mooshammer. 2017. Fluently disfluent? Pauses and repairs of advanced learners and native speakers of German. *International Journal of Learner Corpus Research* 3(2), 118–148. <https://doi.org/10.1075/ijlcr.3.2.02bel>
- Boersma, P. & D. Weenink. 2019. Praat: Doing phonetics by computer (version 6.0.49). <http://www.praat.org/> (accessed 22 March 2019).
- Bosker, H. R., J. Tjiong, H. Quené, T. Sanders & N. D. de Jong. 2015. Both native and non-native disfluencies trigger listeners' attention. In: *DiSS 2015, Proceedings of the 7th Workshop on Disfluency in Spontaneous Speech*, 8-9 August 2015, University of Edinburgh, Scotland, UK.
- Cantero Serena, F. J. 2019. Análisis prosódico del habla: más allá de la melodía [Prosodic speech analysis: beyond the melody]. In: M. R. Álvarez Silva; A. Muñoz Alvarado & L. Ruiz Miyares (eds.), *Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines*. Volumen II, Santiago de Cuba: Ediciones Centro de Lingüística Aplicada, 485–498.
- García-Amaya, L. 2015. A longitudinal study of filled pauses and silent pauses in second language speech. In: *DiSS 2015, Proceedings of the 7th Workshop on Disfluency in Spontaneous Speech*, 8-9 August 2015, University of Edinburgh, Scotland, UK.
- Garrido Almiñana, J. M., Y. Laplaza & C. L. García. 2017. La caracterización pragmática y prosódica de la vocalización “mmm” en español [Pragmatic and prosodic characterization of the vocalization “mmm” in Spanish]. In: V. Marrero Aguiar & E. Estebas Vilaplana (eds.), *Tendencias actuales en fonética experimental: Cruce de disciplinas en el centenario del Manual de Pronunciación Española (Tomás Navarro Tomás)*, Madrid: Uned, 125–129.
- Gósy, M. 2002. A megakadásjelenségek eredete a spontán beszéd tervezési folyamatában [The origin of disfluency phenomena in the spontaneous speech planning process]. *Magyar Nyelvőr* 126, 192–204.
- Horváth, V. 2014. *Hezitációs jelenségek a magyar beszédben* [Hesitation phenomena in Hungarian speech]. *Beszéd. Kutatás. Alkalmazás*. Budapest: ELTE Eötvös Kiadó.
- Medgyes, P. 2001. When the teacher is a non-native speaker. *Teaching pronunciation* 5(12): 429–442.
- Neuberger, T. 2014. *A spontán beszéd sajátosságai gyermekkorban* [Characteristics of spontaneous speech in childhood]. *Beszéd. Kutatás. Alkalmazás*. Budapest: ELTE Eötvös Kiadó.
- Rieger, C. L. 2003. Disfluencies and hesitation strategies in oral L2 tests. In: R. Eklund (ed.), *Proceedings of DiSS'03, Disfluency in Spontaneous Speech Workshop, Gothenburg Papers in Theoretical Linguistics* 90, 5–8 September 2003 2003, Göteborg University, Sweden, 41–44.