SECOND LANGUAGE FLUENCY: RE-THINKING UTTERANCE FLUENCY FROM A PHONETICS-PHONOLOGY INTERFACE

Isadora Reynolds Cavallieri[a], Gillian Wigglesworth[a], Olga Maxwell[a]

[a]The University of Melbourne, Australia
reynoldsi@student.unimelb.edu.au; g.wigglesworth@unimelb.edu.au; omaxwell@unimelb.edu.au

ABSTRACT

Research into second language fluency has called for cross-linguistic studies to rule out measures that can be attributed to intra-speaker variation. However, cross-linguistic comparisons in fluency studies are problematic. Research in the area has not necessarily taken into consideration differences across languages such as syllable structure, phonotactics, durational cues to prominence and prosodic levels, and idiosyncratic nature of pause duration.

The preliminary results of this study into L2 fluency in Chilean Spanish speakers of English revealed that speed and pause phenomena were mostly idiosyncratic, and that segments rather than syllables could be a more reliable measure. Durational cues for phrasal level prominence were not implemented consistently in the L1 and pre-boundary lengthening in the L2 was not necessarily being used to signal prosodic constituent boundaries. It may be useful to re-operationalize measures used in L2 fluency studies from a phonetics-phonology interface perspective.

Keywords: L2 fluency, utterance fluency, speech fluidity, pre-boundary lengthening, pause.

1. INTRODUCTION

Second language fluency (SLF) has been defined as the ability to speak at length with a few pauses [8]. There are two conceptualisations, broad and narrow, of second language fluency which address the need for its more systematic study [23]. The broad definition equates fluency to all-round oral proficiency [16]. The narrow approach views fluency as a temporal phenomenon that can be studied and measured in terms of the properties of the sound wave [6]. In the narrow conceptualization, utterance fluency, also known as speech fluidity, has been studied for over three decades in terms of quantity and quality of speech. Within fluidity, there is a further subdivision into three components: speed, breakdown, and repair fluency. The first corresponds to the speed of speech, the second to the use of silence, and the third to repair mechanisms deployed by speakers to avoid communication breakdown [25].

Studies into SLF have investigated which measures of speech fluidity can best account for what happens in L2 speech. Research has relied on about 23 different measures to account for fluidity [14] [23]. [24] narrowed these down to four core measures of fluency: syllable run (SyllR=mean number of syllables articulated between silent pauses), phonation run [s] (PhonR=mean speech time between silent pauses), syllable duration [ms] (SyllD=mean syllable duration), and silent pause duration [s] (SPauD=mean silent pause duration). [24] discarded articulation rate (SyllAR=syllables articulated per phonation minute) and silent pause rate (SPauR=silent pauses per phonation minute), although these measures have been widely used both in SFL studies and phonological research into speech tempo with somewhat consistent results. Some studies also suggest that SyllAR correlates highly with fluency judgements [6], especially for beginner learners [5]. In terms of breakdown fluency, shorter SPauD has been said to account for more 'fluent' speech [27] with longer SPauD reflecting a higher cognitive load for speech production [26].

In order to decide which measures of SLF can provide the most accurate account of speech fluidity, [23] suggested cross-linguistic analyses where measures of speakers' L1 are contrasted with measures of their L2. However, there are issues with the operationalization of the measurements used in SLF studies. Firstly, two typologically different languages cannot be compared unproblematically using syllables as an underlying unit. This is especially true in the case of English vs Spanish, where the former allows for syllables with several elements while the later favours the CV structure. Hence, comparing articulation rates across languages may render 'inflated' results which may make one language seem inherently faster than the other [22]. Secondly, the use of pauses has been found to be idiosyncratic in different languages and speakers [3] [21], and it seems that SPauD is not as important for fluent L2 speech as is the distribution of these pauses. [11] proposed looking at the effects of dysfluent pauses, that is, those not occurring at constituent boundaries, as this type of pause may be a better reflection of the cognitive load of a task than SPauD. A final issue is that SLF studies have not addressed the effects of language-specific prosodic[1] features on

syllable length, hardly ever looking into how lexical (lexical stress) and post-lexical (accent) prominence can be related to fluency (except for [10] and [15]). One phenomenon which has not been addressed is pre-boundary lengthening (PBL). Segments in a word at constituent boundary tend to be longer than if the same segments occur in phrase-medial position [4]. Lengthening has been found to occur both in the final syllable of an IP and in initial stressed syllable of the word. This seems to be a near universal phenomenon as it is found in most languages [9], but its phonetic implementation (i.e., the amount of lengthening) differs across them [18]. For example, PBL is known be less pronounced in Spanish than in English [18] [19]. Therefore, SyllD and SyllAR scores may be skewed by the effects of PBL on syllable duration due to the difference of implementation across languages. With this in mind, the present study attempts to:

- Determine whether segments rather than syllables are more appropriate as the underlying unit for the operationalization of speech fluidity in cross linguistic studies.
- Explore pause phenomena in Spanish and English.
- Explore the relationship between prosodic durational cues and speech fluidity.

One underlying objective of this study is to incorporate a phonetics-phonology interface into the study of L2 utterance fluency by looking at phonetic implementation of phonological structures. For example, this study acknowledges that the two languages have different segmental inventories and syllable formation processes, and these can impact measures of speech fluidity. For PBL, this study looks at how (and if) it is implemented in both languages and whether this implementation has any effect on SLF. A final objective of the study would be to explore the under-documented L2 English production of Chilean Spanish speakers in terms of SLF.

2. METHODOLOGY

2.1. Participants

The participants were six Chilean Spanish speakers of L2 English (3 females and 3 males, age *M*=32). All of them had an undergraduate degree, they had lived in Australia for an average of 28 months, and their proficiency levels ranged from upper intermediate to advanced[2]. They started learning English during high school as an additional subject, with some training during university. They reported using English in Australia in education and work contexts, mostly in interactions with other L2 speakers.

2.2. Methods

The participants were recorded performing two tasks, both in English and in Spanish. The first was a reading task and the second was a story retelling task where the participants were provided with two stories controlled for cognitive load in terms of the number of characters and plot complexity [17], which they had to narrate in their own words. The audio samples were recorded using the software Samplitude ProX and Charter Oak E700 condenser microphones. These were then processed using WebMAUSBas [13] for automatic segmentation of speech and the alignment of syllables and segments were then manually checked and corrected using the software Praat. The number and duration of segments, syllables, and pauses were extracted using a Praat script to obtain measures of utterance fluency and lengthening. Means, standard deviations and Pearson correlations were calculated using IBM SPSS® Statistics v. 25.

2.1.1. Operationalization of measures

For utterance fluency, the measures chosen were STR, SyllR, PhonR, SyllD, SPauD, SyllAR, and SPauR. The study also proposed three new measures using the same operationalization as SyllAR, SyllD, and SyllR but using segment instead of syllables: segment articulation rate (SegAR=segment per phonation minute), segment duration (SegD=mean segment duration), and segment run (SegR=mean number of segments articulated between silent pauses). These were introduced to explore whether segments were a more appropriate underlying unit than syllables when comparing Spanish and L2 English spoken production.

For pauses, the criteria for cut-off was any silent pause over 100 [ms] which was not an articulatory pause. This choice was made based on [2] who noticed a peak in the presence of brief pauses from 100-150 [ms] both in English and in Spanish. This threshold is much lower than the one usually used in SFL studies, but the cut-off point in previous research has normally been based on researchers' intuitions rather than theory [12, 24]. The study also differentiated between fluent (FlPauN) and dysfluent pauses (DysPauN) in the analysis.

Syllables were divided following syllabification rules for Spanish and English. For example, consonants in coda position in Spanish were placed in the onset of the following syllable as the language has a preference for open syllables. Pre-boundary lengthening was measured at the level of the Intonation Phrase (IP) before a constituent boundary marked by a pause. These measures were divided into two groups: syllables bearing the pitch accent in the

IP (labelled AccPBS), and those placed before the IP boundary, but not bearing the pitch accent (labelled UnPBS). This was done to avoid confounding the effect of accentual prominence on final syllable duration.

3. RESULTS AND DISCUSSION

3.1. Measures for speech fluidity

Table 1 presents a comparison between SyllAR and SegAR, where 'percentage of increase' column (Incr.) shows how much these measures increase when comparing Spanish and L2 English samples. As can be seen, measures in Spanish were always higher suggesting that participants spoke 'faster' in their L1 than the L2. But the differences between Spanish and L2 English were narrower when looking at SegAR instead of SyllAR. This suggest that SyllAR renders 'inflated' results in the L1, which could lead to the interpretation of them being much 'slower' (and therefore, less fluent) in their L2.

Table 1: Syllable articulation and segment articulation rates for both languages in the reading task, presented by speaker (P#) and language.

	Lang.	*SyllAR*	*Incr.*	*SegAR*	*Incr.*
	EN	310.588		751.059	
P1	SP	415.949	25%	940.406	20%
	EN	321.376		802.467	
P2	SP	517.352	38%	1169.666	31%
	EN	305.885		749.790	
P3	SP	459.596	33%	1032.673	27%
	EN	353.102		863.839	
P4	SP	494.110	29%	1135.203	24%
	EN	278.991		693.251	
P5	SP	441.338	37%	1011.513	31%
	EN	245.283		614.702	
P6	SP	396.354	38%	898.068	32%

However, a closer look into measures for SegD reveals that these figures are still not reliable enough for cross-linguistic studies. SegD is always smaller in Spanish (SegD for Spanish in the reading task *M*=.059, *SD*= .006; for the story retelling task *M*=.074, *SD*=.006; for L2 English in the reading task *M*=.081, *SD*=.010; for the story retelling *M*=.091, *SD*=.011) which may be explained by the fact that there is no phonemic length distinction for vowels in the language. Though SegD may not be the most accurate measure, segments might still be a more valid unit when comparing English and Spanish than syllables. Pearson correlations between variables of speech fluidity and STR[3] reveal that the measures created for this study using segments are highly and significantly correlated to STR, where the correlations were more robust than for measures using syllables (SyllAR R_S=.416, p=.043 vs. SegAR R_S=.487, p=.016; SyllR R_S =.662, p=.001 vs. SegR R_S=.684, p <.001).

3.2. Pause duration and distribution

When comparing the use of pauses, duration and distribution seemed to be related to the type of task being performed[4]. Pauses in Spanish were longer than in L2 English in story retelling and very similar in length in the reading task, as can be seen in Table 2. These results seem counterintuitive considering that story retelling is more cognitively demanding than reading. It could be concluded that speech style has a considerable effect on the use of pauses and that style is idiosyncratic to the languages under study.

Table 2: Number of silent fluent and dysfluent pauses, and means of silent pause duration presented by task (reading task (RT), story retelling (SR)) and language.

		FlPauN		*DysPauN*		*SPauD*	
Task	Lang	*M*	(SD)	*M*	(SD)	*M*	(SD)
RT	EN	25.83	11.14	5.33	4.76	0.35	0.05
	SP	15.33	1.75	0.33	0.52	0.34	0.05
SR	EN	31.16	6.49	8.50	2.35	0.55	0.12
	SP	23.00	5.93	3.67	2.07	0.70	0.12

The correlations between SPauD and type of task[5] in the L1 (R_S=-.908, p < .001) and in the L2 (R_S=-.782, p=.003) show that pauses in this sample were shorter in the reading task but this difference in length was more significant in Spanish. These results seem to reinforce the fact that speaking style is more relevant than cognitive load when it comes to pause duration. Conversely, the number of pauses produced in the different tasks, especially DysPauN, may be a better indicator of the cognitive load of the task than SPauD, given a more frequent occurrence of these in English and even more so in the story retelling task. Added to this, several dysfluent pauses in were found in the 100-250 [ms] (i.e., brief pause) range, pointing to the need of including brief pauses in SLF studies.

3.2. Pre-boundary lengthening

In the reading task[6] in English, AccPBS were the longest for all speakers (*M*=.292, *SD*=.059), while UnPBS were the shortest (*M*=.153, *SD*=.027) even when compared to SyllD (*M*=.201, *SD*=.026). A

possible conclusion is that participants can imitate prosodic features of the L2 implementing length as a marker for post-lexical accent in pitch accented syllables but not for PBL in syllables at IP boundary, even though this is a salient feature in L1 English. In Spanish, length was not consistently implemented neither for marking post-lexical accent nor to signal constituent boundaries. This is in line with previous studies where duration is not systematically used in Spanish to cue post-lexical accent or PBL [1] [20]. Figures 1 and 2 show the results for two participants with the lowest (P4) and the highest (P6) proficiency. AccPBS were longer than UnPBS and SyllD for both speakers in L2 English. These figures illustrate the use of length as a cue to post-lexical accent more than as a marker of IP boundaries. The results for Spanish show high variability between speakers, pointing to the aforementioned lack of consistency in the implementation of length[7]

Regarding the presence of PBL as a marker of speech fluidity, in the L2 samples AccPBS length highly correlated with PhonR (R_S=-.716, p=.009), which could mean that longer runs result in longer accented syllables. UnPBS, on the other hand, correlated highly with SPauD (R_S=.807, p=.002), meaning that syllables are longer before longer pauses. These results suggest that there is a relationship between syllable lengthening and speed and breakdown fluency, and this cannot really be explained by PBL in this sample. Moreover, in Spanish AccPBS and UnPBS (and even SyllD) were negatively correlated to both SyllAR and SegAR, meaning that syllable lengthening in Spanish is susceptible to speed more than any other factor, reinforcing the idea that there is something in final lengthening in L2 English which is specific to this language and may, in some way, be related to speech fluidity.

Figure 1: Clustered boxplot for AccPBS, UnPBS and SyllD (in seconds) for participant 4, reading task in English (P4(E)) and Spanish (P4(S)).

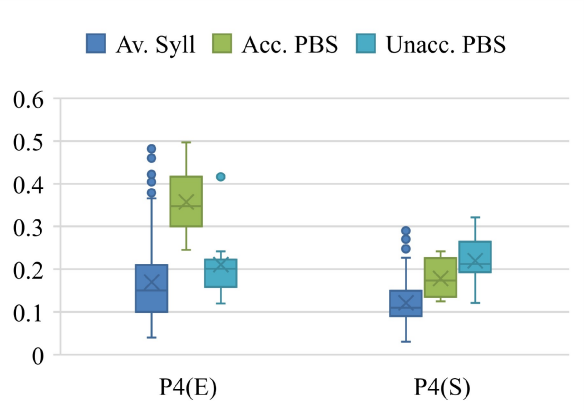
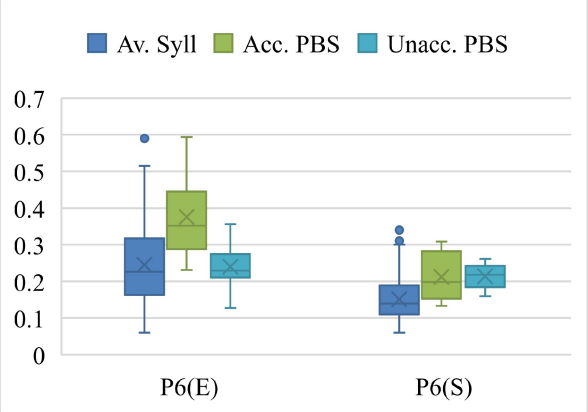


Figure 2: Clustered boxplot for AccPBS, UnPBS and SyllD (in seconds) for participant 6, in the reading task in English (P4(6)) and Spanish (P6(S)).



4. CONCLUSIONS

This study proposes a revision of the measures used to measure utterance fluency (speech fluidity) in SLF studies. This cross-linguistic analysis relied on the phonetics-phonology interface, looking at what makes the two languages different at the phonological level, how this is implemented phonetically, and whether this implementation has any effect on measures of speech fluidity. From a closer look into syllables in both languages it was evident that this unit of analysis was problematic. Although far from conclusive, the study does offer evidence to call for a revision of the operationalization of speech fluidity measures.

SLF in terms of speed seems to be dependent on the speakers' style, but it is hard to find unequivocal support for this with the current range of measures. In the study, pause length was mostly idiosyncratic. It may be that the number of dysfluent pauses could reveal more about SLF than the length of these. Finally, syllable lengthening appears to relate to measures of speed and breakdown fluency in L2 English, although it appears that this may be influenced by some other factor than the marking of prosodic boundaries Despite the small sample size, these findings emphasise the need to revisit the way in which speech fluidity is measured. The next steps in this study are to include a larger number of participants, as well as to conduct further work to find measures that can better account for temporal phenomena in terms of speed, pause, and duration in SLF. This is especially relevant for cross-linguistic studies if researchers are to identify which measures account exclusively for phenomena in L1 speech.

5. REFERENCES

[1] Alcoba, S., & Murillo, J. (1998). Intonation in Spanish. In Hirst, D., Cristo, A. D. (eds), *Intonation systems: a survey of twenty languages* (pp. 152-166): Cambridge University Press.

[2] Campione, E., Véronis, J. 2002. *A large-scale multilingual study of silent pause duration. Proc. Speech Prosody,* Aix-en-Provence.

[3] Cenoz, J. 2000. Pauses and hesitation phenomena in second language production. *ITL-International Journal of Applied Linguistics* 127(1), 53-69.

[4] Cho, T., Kim, J., Kim, S. (2013). Preboundary lengthening and preaccentual shortening across syllables in a trisyllabic word in English. *The Journal of the Acoustical Society of America*, *133*(5), 384-390.

[5] Cucchiarini, C., Strik, H., Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *J. Acoust. Soc. Am*. 111(6), 2862-2873.

[6] De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R., Hulstijn, J. H. 2013. Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics* 34(5), 893-916.

[7] Derwing, T. M., Munro, M. J., Thomson, R. I., Rossiter, M. J. 2009. The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition* 31(4), 533-557.

[8] Fillmore, C. J. (1979). On fluency. In: Fillmore, C. J., Kempler, D., Wang, W. S.-Y. (eds), *Individual differences in language ability and language behavior* (pp. 85-101). New York: Academic Press.

[9] Fletcher, J. 2010. The prosody of speech: Timing and rhythm. In: Hardcastle, W., Laver, J. (eds), *The Handbook of Phonetic Science*. Oxford: Blackwell, 521-602.

[10] Hieke, A. E. 1985. A componential approach to oral fluency evaluation. *The Modern Language Journal* 69(2), 135-142.

[11] Kahng, J. 2018. The effect of pause location on perceived fluency. *Applied Psycholinguistics* 39(3), 569-591.

[12] Kirsner, K., Dunn, J., Hird, K., Parkin, T., Clark, C. 2002. *Time for a pause. Proc. 9thASST* Melbourne, 52-57.

[13] Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language, 45*, 326-347.

[14] Kormos, J. 2006. *Speech Production and SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.

[15] Kormos, J., Dénes, M. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32(2), 145-164.

[16] Lennon, P. 2000. The lexical element in spoken second language fluency. In: Riggenbach, H. (ed), *Perspectives on fluency* Ann Arbor: University of Michigan, 25-42.

[17] Luoma, S. 2004. *Assessing speaking*. Cambridge: Cambridge University Press.

[18] Ortega-Llebaria, M., Prieto, P. 2007. Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables. In: Prieto, P., Mascaró, J., Solé, M.J. (eds), *Segmental and prosodic issues in Romance phonology*. Amsterdam: John Benjamins Publishing, 155-176.

[19] Prieto, P., del Mar Vanrell, M., Astruc, L., Payne, E., Post, B. 2012. Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication* 54(6), 681-702.

[20] Puebla, J. I., Escudero Mancebo, D. 2014. Alargamientos y pausas dentro de palabra como rasgo prosódico no lingüístico del español no peninsular: estudio preliminar. *Proc. 5th Congreso de Fonética Experimental*, 1-15.

[21] Riazantseva, A. 2001. Second language proficiency and pausing a study of Russian speakers of English. *Studies in SLA* 23(4), 497-526.

[22] Roach, P. 1998. Some languages are spoken more quickly than others. In: Bauer, L., Trudgill, P. (eds), *Language Myths*. Harmondsworth: Penguin 150–158.

[23] Segalowitz, N. 2010. *Cognitive bases of second language fluency*. New York, NY: Routledge.

[24] Segalowitz, N., French, L., Guay, J.-D. 2018. What Features Best Characterize Adult Second Language Utterance Fluency and What Do They Reveal About Fluency Gains in Short-Term Immersion? *Canadian Journal of Applied Linguistics* 20(2), 90-116.

[25] Tavakolil, P., Skehan, P. 2005. Strategic planning, task structure, and performance testing. In: Ellis, R. (ed), *Planning and task performance in a second language*. Amsterdam: John Benjamins, 239-276.

[26] Towell, R. 2002. Relative degrees of fluency: A comparative case study of advanced learners of French. *IRAL* 40(2), 117-150.

[27] Vercellotti, M. L. 2012. *Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other.* PhD dissertation, University of Pittsburgh.

1 Prosody here refers to patterns of lexical and post-lexical prominence, pitch, rhythm, and intonation in utterances.

2 These participants had to have obtained at least an equivalent of 6.5 in the IELTS© to enter the country.

3 Some studies have used STR as a measure with which to correlate others because it is supposed to account for proficiency (or broad fluency) more than utterance fluency.

4 For the purpose of this analysis, task type is equated to speaking style being used, i.e., reading vs telling a story.

5 For these calculations, the reading task was used as a dummy variable holding story retelling as the constant.

6 The story narration will not be discussed due to space constraints, but results were similar across tasks.

7 The irregular results for UnPBS for P4 (which were the longest) can be explained by the nature of the syllables at boundary position. Due to a paroxytonic tendency of Spanish, most syllables in this position were unaccented but these mostly contained /s/ or the aspirated version of /s/ in coda position, which were lengthened by this speaker.