

Prosodic Cues to Monolingual versus Code-switching Sentences in English and Spanish

Page Elizabeth Piccinini¹, Marc Garellek¹

¹Department of Linguistics, University of California, San Diego, La Jolla, California, USA

ppiccinini@ucsd.edu, mgarellek@ucsd.edu

Abstract

Code-switching offers an interesting methodology to examine what happens when two linguistic systems come into contact. In the present study, two experiments were conducted to see if (1) listeners were able to anticipate code-switches in speech-in-noise, and (2) prosodic cues were present in the signal as potential cues to an upcoming code-switch. A speech-in-noise perception experiment with early Spanish-English bilinguals found that listeners were indeed able to accurately identify words in code-switching sentences with the same accuracy as in monolingual sentences, even in highly-degraded listening conditions. We then analyzed the stimuli used in the perception experiment, and found that the speaker used different prosodic contours for code-switching productions compared to monolingual productions. We propose that listeners use specific code-switching prosody to anticipate code-switches, and thus ease processing costs in word identification.

Index Terms: prosodic contours, bilingualism, code-switching

1. Introduction

Code-switching – where two languages are used in a single utterance, sometimes switching mid sentences – is a common practice among high-proficiency bilinguals [1]. Code-switching is well documented with regards to sociolinguistic practices [2, 3, 4] and its syntactic and morphological aspects for several language pairings [5, 6, 7, 8, 9].

Psycholinguists have also studied code-switching, by examining how bilinguals deal with processing one language directly after another. Most switching studies have used an artificial form of switching, namely picture naming – where bilinguals had to switch the language in which they named a picture. [10, 11, 12, among others]. One consistent result across these studies is that bilinguals take longer to name pictures in switch trials compared to non-switch trials, regardless of the direction of the switch. A cost has also been found for processing code-switches in electrophysiological studies, both for single lexical items [13] and when comparing utterances that are either monolingual or code-switching [14, 15]. Given these costs, it seems odd that code-switching would be such a productive practice for bilinguals.

However, most of this work examined processing of words in isolation. Further, in studies that did provide full sentences, the stimuli were presented orthographically. Yet, it is possible that acoustic correlates to code-switches exist in the speech stream, and that listeners use them as cues to an upcoming switch; these cues may then reduce some of these processing costs. For example, it is well documented that English and Spanish, besides having different phonetic inventories, also have different intonation patterns. Notably, English declaratives

typically have shallow-rise prenuclear and nuclear pitch accents (e.g., H* in the ToBI labeling conventions), whereas (Mexican) Spanish uses scooping rises with delayed peaks (L+>H*) in prenuclear position, but usually low-f0 targets (L*) as its nuclear pitch accent. [16, 17, 18]. In code-switching, it is possible that speakers will produce intonation patterns that do not directly map on to either language. Listeners could use this information to their advantage to anticipate upcoming language switches. For example, work on Spanish-English code-switching and narrow focus has found that while code-switching cues narrow focus, it is not enough for the sentence to include a code-switch; the sentence must also have the correct prosody for the sentence to be consistently understood as signaling narrow focus [19].

The present study asks two main questions. First, do Spanish-English bilinguals have trouble identifying words in code-switches sentences because code-switching induces high processing demands? To examine this question, we conducted a speech-in-noise study using both monolingual and code-switching utterances. If there truly is a processing deficit due to code-switches' being unexpected, then listeners should perform worse on the code-switching sentences compared to the monolingual sentences. However, if they perform equally well on code-switching and monolingual sentences, then it is possible that there are phonetic cues that allow listeners to anticipate a code-switch. As already discussed, one known difference between English and Spanish is pitch accent realization. Thus the second question of the study is, are there systematic prosodic differences between English, Spanish, and code-switching (both English to Spanish and Spanish to English) sentences? This will be investigated by examining the F0 contours of the full sentences used as stimuli in the speech-in-noise study, as well as the F0 realization of the pitch-accented words in each utterance.

2. Part 1: Perception experiment

2.1. Methods

2.1.1. Listeners

Eight early Spanish-English bilinguals (seven female, one male) of Mexican-American heritage participated in the perception experiment. The average age was 19.6 (standard deviation 1 year). Average age of acquisition of Spanish was 1.3 years (standard deviation 1.1 years); that of English was 3.1 years (standard deviation 1.8 years). We administered the Bilingual Dominance Scale [20] to all participants. The scale computes a score for how dominant the participant is in one language or another: a score of 0 means that the participant is a balanced bilingual; a high positive score mean the participant is heavily

English dominant; a high negative score means heavily Spanish dominant. The average score for the participants was 6.9 (standard deviation 7.7), suggesting that listeners were English dominant, despite the fact that their first language was Spanish. All participants were undergraduates at the University of California, San Diego, and received course credit in exchange for their participation in the study.

2.1.2. Stimuli

Sentences were modified versions of the Bamford-Kowal-Bench (BKB) sentences [21], which have been used in several past speech-in-noise studies [22, 23, 24]. The sentences were produced by a bilingual female speaker of American English and Mexican Spanish for use as English, Spanish, and code-switching sentences (code-switching English to Spanish (henceforth, **CS-ES**) and code-switching Spanish to English (**CS-SE**)). More information about the speaker's language background is provided in 3.1.1. Sentences were translated for use in the Spanish and code-switching sentences. The original BKB sentence list was created to include short sentences with vocabulary commonly used by partially-hearing children ages eight to 15. Similarly, the Spanish translations were modified as needed to include high-frequency words in Spanish, as well as be culturally appropriate for the dialect of Spanish spoken by the bilingual participants (viz., Mexican Spanish). Similar sentence structures were maintained for both the Spanish translations and the code-switching sentences. Code-switch location within a sentence was counterbalanced for syntactic position. This was done to reflect the grammatical rules of code-switching as discussed in previous research on Spanish-English code-switching [5, 6, 7]. A native Spanish-English bilingual reviewed the code-switching sentences for grammaticality and naturalness. The BKB sentences are blocked into lists of 16 sentences, each containing 50 keywords. Keywords were salient words in the sentence, generally nouns and verbs. Eight lists were used for the experiment, and half of an additional list was used for an initial practice session. Thus, there was a total of 32 test items per language context. Sentences were mixed with white noise at four signal-to-noise ratios (SNRs), -6, -3, +0, and +3 dB in Praat. There were two lists per SNR.

2.1.3. Procedure

Listeners completed a speech-in-noise task, during which they wrote down as many words as possible for each sentence presented. Listeners began with a practice session consisting of eight sentences (two from each language context) with feedback. After the practice session, listeners were presented with two blocks of four lists of sentences, one list per SNR per block. Within a given list, there was an equal number of sentences from each language context (four English, four Spanish, four CS-ES, four CS-SE). Stimuli were randomized within list and the list order was randomized within block.

2.2. Results

Regarding SNR, listeners do well in all four contexts, and perform above chance (50% correct) even at the lowest SNR of -6 dB. A logistic mixed-effects regression model was run to test for the effect of SNR on keyword identification. Keyword identification ('correct' or 'incorrect') was the dependent variable, SNR was the independent variable, and participant was included as a random slope by SNR. As expected, listeners identified words significantly better as the SNR increased [-6 vs -3: β

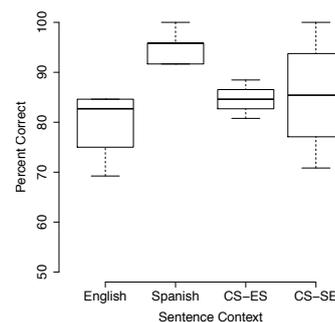


Figure 1: Percent correct for keywords identified in each context at an SNR of -6 dB (the 'high noise' context).

$= 0.26, z = 1.66, p = 0.10$; -6 vs +0: $\beta = 0.84, z = 4.66, p < 0.001$; -6 vs +3: $\beta = 1.81, z = 6.32, p < 0.001$). Since listeners were largely at ceiling for the higher SNR levels, a model was also run on the -6 SNR ('high noise') data alone, in order to examine more closely the effects of language context. Keyword identification ('correct' or 'incorrect') was the dependent variable, language context was the independent variable, and participant was included as a random slope by language context. These results are presented in Figure 1. Listeners were significantly better at Spanish than the three other language contexts [Spanish vs English: $\beta = -1.53, z = -4.16, p < 0.001$; Spanish vs CS-ES: $\beta = -1.12, z = -3.17, p < 0.01$; Spanish vs CS-SE: $\beta = -1.03, z = -2.42, p < 0.05$]; there were no significant differences among any of the other contexts.

2.3. Interim discussion

The results of the perception experiment demonstrate that listeners are good at identifying words in code-switching sentences, even in listening conditions with high noise. Listeners performed best at the Spanish context. This is an interesting finding given that they are English dominant. However, this may be an effect of Spanish being their L1, thus demonstrating that current dominance is not always the most important factor. Listeners performed equally well on code-switching sentences as English sentences (their dominant language). This result seems to run counter to previous research which found a processing deficit in code-switching. This may be because previous experiments used only orthographic or pictorial switches and did not provide audio. If indeed phonetic cues in the signal can be used by listeners to anticipate a code-switch, then it is understandable that processing code-switches becomes difficult in the absence of audio information, which should result in longer processing times. It should also be noted that past studies that found a processing deficit for code-switching tested online processing, while the present study does not. However, this study does provide at least initial data suggesting that there is *not* a processing deficit for code-switching. To test if listeners in Part 1 performed well on code-switching stimuli because of potential phonetic cues to code-switching, in Part 2 we examine the prosodic structure of the stimuli. We hypothesize that prosodic differences – namely, differences in F0 contours over the utterance and in pitch accent realizations – may be present in code-switching contexts.

3. Part 2: F0 analysis of stimuli in Part 1

3.1. Methods

3.1.1. Speaker

One speaker produced all stimuli. She is a 22 year-old early Spanish-English bilingual having learned first Spanish at home and then English around age 6. She is a speaker of both Mexican and Peninsular Spanish, although she was instructed to produce the Spanish and code-switching sentences with a Mexican Spanish accent. She is a native speaker of Southern Californian English. She had a score of 1 on the Bilingual Dominance Scale suggesting she is a balanced bilingual. At the time of recording, she was naive to goals of the experiment. A larger-scale analysis with more bilingual speakers is underway.

3.1.2. Stimuli

Sentences were the test stimuli used in Experiment 1.

3.1.3. Method of analysis

For each sentence type (English, Spanish, CS-ES, and CS-SE), the F0 contours were analyzed both across the entire sentence and at specific points in the sentence. For the entire sentence, F0 values were extracted at 1% increments, starting at 0% into the sentence up to 100% into the sentence, resulting in a total of 101 measurements per sentence. Because content words always bore pitch accents, we also segmented each content word's stressed vowel and any adjacent sonorants to the stressed vowel within the syllable. Each sentence was uttered as one Intonation Phrase, with no IP-medial intermediate phrase break. F0 values were taken at 5% increments, starting at 0% into the stressed syllable up to 100%, resulting in a total of 21 measurements per stressed syllable. Syllables were coded as being the first stressed syllable (and thus, first pitch accent) in the Intonation Phrase, an Intonation Phrase-medial stressed syllable (sometimes there were up to three medial stressed syllables for a given Intonation Phrase), or the final stressed syllable (i.e., the nuclear-pitch-accent syllable) in an intermediate phrase (and thus, also the Intonation Phrase). Once the F0 values were obtained, outliers (based on visual inspection of the F0 tracks) were manually removed.

3.2. Results

The F0 contours across the entire sentence for the four contexts are shown in Figure 2. Visual inspection suggests that there are only two contours in use. One contour type, with a more extreme but delayed initial rise, is used in the English and CS-SE contexts. A second contour type, with a shallower and earlier initial rise, is used in the Spanish and CS-ES contexts. Towards the end of the code-switching sentences, the F0 generally maps onto the target language for the end of the sentence (i.e. CS-ES maps onto the contour for the Spanish sentence), with some deviation.

Figures 3, 4, and 5 present the contours for the first, middle, and final stressed syllables (which were also pitch-accented) in each context. To test for differences between contours, ANOVAs were conducted at 0%, 50%, and 100% for each syllable grouping. The dependent variable was F0 (in Hz) at the given time point, and the independent variable was language context. For the first stressed syllables in the sentence, there were significant effects of context at all three time points [0%: $F(3, 110) = 3.29, p < 0.05$; 50%: $F(3, 120) = 6.22, p < 0.001$; 100%: $F(3, 121) = 9.39, p < 0.001$]. Post-hoc Tukey tests were

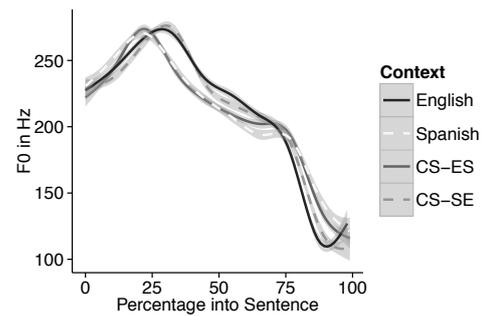


Figure 2: Normalized F0 contour for full sentence by context. Error bars indicate one standard error.

conducted to determine which specific contexts differed significantly from each other. English and Spanish pitch-accented vowels significantly differed from each other at the vowel's midpoint [$p < 0.001$] and end [$p < 0.001$], with F0 being higher at both time points in the English context. The F0 of CS-ES was significantly lower than that of English at the beginning of pitch-accented vowels [$p < 0.05$], but was significantly higher than that of Spanish at the vowels' end points [$p < 0.001$]. CS-SE had a significantly lower F0 than that of English midway through the vowels [$p < 0.05$].

For the middle stressed syllables, there were significant effects of context at all three time points [0%: $F(3, 176) = 7.35, p < 0.001$; 50%: $F(3, 188) = 5.34, p < 0.01$; $F(3, 184) = 9.03, p < 0.001$]. Post-hoc Tukey tests were conducted to check which contexts differed from each other. English and Spanish significantly differed from each other at all three time points [0%: $p < 0.001$; 50%: $p < 0.001$; 100%: $p < 0.001$], with English having a consistently higher F0 than Spanish. CS-ES had a significantly lower F0 than English at the beginning of the vowels [$p < 0.05$], and a significantly higher F0 than Spanish at the end points [$p < 0.05$]. CS-SE also significantly differed from English at the vowel end points [$p < 0.01$], where it had a lower F0 than English.

For the final stressed syllables, there were significant effects of context at 0% and 100% [0%: $F(3, 105) = 4.32, p < 0.01$; 100%: $F(3, 99) = 7.65, p < 0.001$]. Post-hoc Tukey tests were conducted to determine which contexts differed from each other. English had a significantly higher F0 than Spanish at both beginnings and ends of vowels [0%: $p < 0.05$; 100%: $p < 0.001$]. CS-ES significantly differed from English at the end points [$p < 0.05$], where it had a lower F0 than English. CS-SE has a significantly higher F0 than Spanish at the beginning and end points [0%: $p < 0.05$; 100%: $p < 0.01$].

In sum, the stressed (and pitch-accented) syllables in the utterance differed depending on the language context. English stressed syllables generally had a higher F0 than Spanish ones. On the other hand, stressed syllables in the code-switching contexts usually had an intermediate F0 – not as high as English, but higher than Spanish.

4. General discussion

The goals of the present study were to see if (1) listeners are able to anticipate code-switches in speech-in-noise, and (2) prosodic cues are present in the signal to warn of an upcoming code-switch. The results of the perception experiment suggest

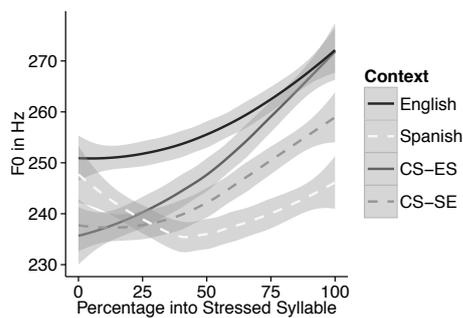


Figure 3: Normalized F0 contour for first stressed syllable in sentence by context. Error bars indicate one standard error.

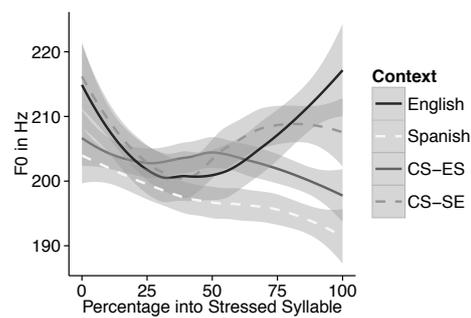


Figure 5: Normalized F0 contour for final stressed syllable in sentence by context. Error bars indicate one standard error.

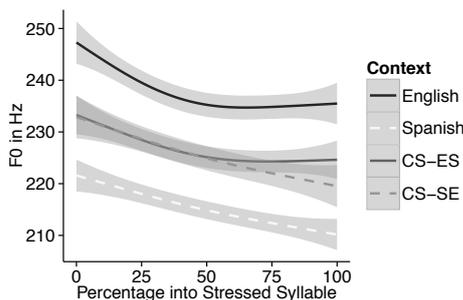


Figure 4: Normalized F0 contour for middle stressed syllables in sentence by context. Error bars indicate one standard error.

that listeners are able to easily process code-switches, even in a detrimental environment such as high noise. Listeners performed equally well at code-switches compared to monolingual sentences in English. This suggests that there could be something in the signal that cues listeners of an upcoming switch, which eases processing costs later on. One possibility is that listeners use syntactic and semantic information to anticipate a code-switch. This conclusion is unlikely here, as all sentences were derived from English monolingual sentences and thus could be completed in either language. Conversely, there may well be phonetic cues, either segmental or suprasegmental. The aim of Part 2 was to determine if there were prosodic cues (in the form of F0 differences) that listeners may use to anticipate a code-switch.

The results of Part 2 found that the speaker who produced the stimuli in Part 1 did in fact produce different F0 contours for code-switching sentences compared to monolingual sentences. English stressed syllables had an overall higher F0 than Spanish ones, consistent with the fact that Spanish has low nuclear pitch accents and scooping rises (with delayed high F0) in pre-nuclear position [17, 18]. On the other hand, English pitch accents tend to be high-toned, even in the nuclear-pitch-accented position [16]. However, in the English-to-Spanish code-switch, the speaker does not simply use English-like tone realization for the first half of the sentence and then immediately switch to Spanish-like tones for the second half of the sentence. Instead, the speaker used intonational patterns from both languages throughout her code-switching sentences. When examining the contours for the entire sentence, the speaker produced the beginning of code-switching sentences with F0 patterns that

were more similar to the post-switch language: e.g., English-to-Spanish code-switching sentences began with more Spanish-like intonation, rather with English-like intonation that became progressively more Spanish-like. This may serve as a cue to the listener of an impending code-switch. When stressed syllables of content words were examined, results were generally mixed: the speaker produced code-switching contours that never fully mapped on to either language. This can be seen most clearly in Figure 4. This is in contrast to the monolingual contours, which were consistently different from each other at almost all time points analyzed. Thus, while the speaker maintains different prosodic patterns for English and Spanish, this does not necessarily map directly onto her code-switching contours. Instead, code-switching F0 contours are in a category of their own, regardless of the direction of the switch (English-to-Spanish vs Spanish-to-English). However, another possible explanation for the differences in F0 is the segmental make-up of the words in each context, particularly the segmental make-up of stressed syllables, which could affect the timing and maximum of F0 peaks. To address this possibility, future analyses will include a phonological analysis of the sentences, to see how contexts differ in terms of tone realization and alignment. Future work will also expand upon the production study by recording additional speakers, to see if the effects shown here hold for the population at large.

5. Conclusion

The present study demonstrated that listeners are able to anticipate code-switches and thus ease processing costs in speech-in-noise word identification. An analysis of the stimuli suggested that this anticipation may in part be due to the F0 contours of the sentences as a whole, as well as those of pitch-accented words, which differ in code-switching contexts compared to monolingual contexts. Additional work on the intonation patterns of naturalistic code-switching, both in production and perception, will test whether code-switching utterances generally differ prosodically from monolingual ones in the ways found here, and whether these differences are used as cues in real-life language processing.

6. Acknowledgements

We thank our speaker for volunteering to record the stimuli and the members of the Phonetics Lab for comments and advice on this project.

7. References

- [1] Gumperz, J. J., "The sociolinguistic significance of conversational code-switching", *RELC Journal*, 8(2):1–34, 1977.
- [2] Milroy, L., and Gordon, M., "Style-shifting and code-switching", in *Sociolinguistics: Method and Interpretation*, 198222, Blackwell Publishing Ltd, 2003.
- [3] Woolard, K. A., "Codeswitching", in A. Duranti [Ed], *A Companion to Linguistic Anthropology*, 73–94, Blackwell Publishing Ltd, 2004.
- [4] Lipski, J. M., "Code-switching or borrowing? No sé so no puedo decir, you know", in L. Sayahi and M. Westmoreland [Eds], *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*, 1–15, Cascadilla Proceedings Project, 2005.
- [5] Pfaff, C. W., "Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English", *Language*, 55(2):291–318, 1979.
- [6] Poplack, P., "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching", *Linguistics*, 18(7-8):581–618, 1980.
- [7] Poplack, P., "Contrasting patterns of code-switching in two communities", in E. Wande, J. Anward, B. Nordberg, L. Steensland and M. Thelander [Eds], *Aspects of Multilingualism*, 51–77, Uppsala, 1987.
- [8] Myers-Scotton, C., "Language contact: Why outsider system morphemes resist transfer", *Journal of Language Contact - TEMA*, 2:21–41, 2008.
- [9] Poplack, S., Zentz, L., and Dion, N., "Phrase-final prepositions in Quebec French: An empirical study of contact, code-switching and resistance to convergence", *Bilingualism: Language and Cognition*, 15(2):203–225, 2011.
- [10] Meuter, R. F. I., and Allport, A., "Bilingual language switching in naming: Asymmetrical costs of language selection", *Journal of Memory and Language*, 40(1):25–40, 1999.
- [11] Costa, A., and Santesteban, M., "Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners", *Journal of Memory and Language*, 50(4):491–511, 2004.
- [12] Schwieter, J. W., and Sunderman, G., "Language switching in bilingual speech production: In search of the language-specific selection mechanism", *The Mental Lexicon*, 3(2):214–238, 2008.
- [13] Chauncey, K., Grainger, J., and Holcomb, P. J., "Code-switching effects in bilingual word recognition: a masked priming study with event-related potentials", *Brain and Language*, 105(3):161–74, 2008.
- [14] Moreno, E. M., Federmeier, K. D., and Kutas, M., "Switching languages, switching palabras (words): An electrophysiological study of code switching", *Brain and Language*, 80(2):188–207, 2002.
- [15] Proverbio, A. M., Leoni, G., and Zani, A., "Language switching mechanisms in simultaneous interpreters: An ERP study", *Neuropsychologia*, 42(12):1636–1656, 2004.
- [16] Beckman, M. E., and Elam, G. A., "Guidelines for ToBI labeling", The Ohio State University Research Foundation, 2007.
- [17] Beckman, M. E., Díaz-Campos, M., McGory, J. T. and Morgan, T. A., "Intonation across Spanish, in the Tones and Break Indices framework", *Probus*, 14:9–36, 2002.
- [18] De-la-Mota, C., Butragueño, M., and Prieto, P., "Mexican Spanish Intonation", in P. Prieto and P. Roseano [Eds], *Transcription of Intonation of the Spanish Language*, 319–350, Lincom Europa, 2010.
- [19] Olson, D., and Ortego-Llebaria, M., "The perceptual relevance of code switching and intonation in creating narrow focus", in M. Ortego-Llebaria [Ed], *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, Cascadilla Proceedings Project, 57–68, 2010.
- [20] Dunn, A. L., and Fox Tree, J. E., "A quick, gradient Bilingual Dominance Scale", *Bilingualism: Language and Cognition*, 12(3):273–289, 2009.
- [21] Bench, J., Kowal, A., and Bamford, J., "The BKB (Bamford-Kowal-Bench) sentence lists for partially hearing children", *British Journal of Audiology*, 13(3):108–112, 1979.
- [22] Bent, T., and Bradlow, A. R., "The interlanguage speech intelligibility benefit", *Journal of the Acoustical Society of America*, 114(3):1600–1610, 2003.
- [23] Bradlow, A. R., Kraus, N., and Hayes, E., "Speaking clearly for children with learning disabilities: Sentence perception in noise", *Journal of Speech, Language, and Hearing Research*, 46:80–97, 2003.
- [24] Van Engen, Kristin J., "Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble", *Speech Communication*, 52:943–953, 2010.