Covert contrast in children's spectral moments' realizations of alveolar and alveopalatal fricatives in French

Margaret Kehoe[1], Marie Philippart de Foy[2]

[1]Université de Genève, [2]Aix-Marseille Université

margaret.winkler-kehoe@unige.ch, marie.phillippart-de-foy@univ-tlse2.f

Abstract. Alveopalatal fricatives /ʃ, ʒ/ are one of the last sets of sounds to be acquired by French-speaking children; they are often substituted by alveolar fricatives [s, z]. Studies that have employed spectral moments analyses show that spectral differences between alveolar and alveopalatal fricatives are not well defined in English-speaking three-year-olds but become so in five-year-olds. Nevertheless, children who are transcribed as making errors (e.g., /ʃ/ transcribed as [s]), may display evidence of a covert acoustic contrast. We investigate the production of alveolar and alveopalatal fricatives by French-speaking children using a spectral moments' analysis. We examine whether children transcribed as making errors are making covert acoustic contrasts in their speech. French-speaking monolingual and bilingual children (n=80), aged 2;6 to 6;10, participated in a naming task in which they produced words containing fricatives /s, ʃ/ in different word positions. Their productions were phonetically transcribed and acoustically analyzed. We measured the spectral mean or centroid for all fricative productions. Data were analyzed using mixed-effects models. We found a significant effect of age, gender, and sound on spectral means. Spectral means decreased with age, were higher for girls than boys, and were higher for /s/ than /ʃ/. Importantly, we documented a significant interaction between /ʃ/ accuracy (based on phonetic transcription) and spectral mean realization. As /ʃ/ accuracy increased, the contrast between centroid values for /s/ and /ʃ/ increased. The difference in centroid values was not significant for low /ʃ/ producers (less than 20% accuracy) but it was for mid (33-67% accuracy) and high /ʃ/ producers (greater than 75% accuracy). The findings suggest that many children display a reliable acoustic contrast between [s] and [ʃ] even though their transcription data indicate that the contrast has not yet been mastered.

Keywords: phonological acquisition; spectral moments; alveopalatal fricatives; acoustic analysis; French-speaking children

Introduction

Alveopalatal fricatives are difficult to produce. In French, they are one of the last sets of sounds to be acquired and are among the sounds most frequently targeted in speech sound intervention (Aicart-de-Falco & Vion, 1987). In this study, we investigate the production of the alveopalatal fricative /ʃ/ by French-speaking children, aged 2;6 to 6;10, and contrast its production with that of the alveolar fricative /s/. We conduct a spectral moments' analysis which details the spectral characteristics of segments in terms of multiple statistical moments. Previous studies in English indicate that children who are transcribed as making errors with alveopalatal fricatives (e.g., /ʃ/ transcribed as [s]), may display evidence of a covert acoustic contrast (Li et al., 2009). We

investigate whether covert contrasts are also present in the speech of French-speaking children acquiring alveopalatal fricatives.

Acquisition of alveolar and alveopalatal fricatives

Fricatives are sounds produced by air passing through a narrow constriction in the vocal tract. Alveopalatal, as opposed to alveolar fricatives, require motor differentiation of the blade versus tongue tip: /ʃ/ has a wider tongue groove resulting in a larger cross-sectional area than /s/, and the constriction for /ʃ/ is further back in the vocal tract than for /s/.

In French, MacLeod et al. (2011) report that /s, ʃ/ are late sounds being acquired after 4;5. Similarly, Aicart-de-Falco and Vion (1987), in a study of European French-speaking children, aged 3 to 6 years, report late acquisition of alveolar and alveopalatal fricatives. They note that over 60% of all consonant errors concern the sounds /s, ʃ/ (as well as /z, ʒ/) and that errors continue through to six years.

Covert contrast

When the acquisition of a sound contrast is protracted, children may pass through a stage in which they make a statistically significant acoustic difference between two sounds but one that is not perceptually reliable. This is referred to as a covert contrast. A well-known example is Macken and Barton's (1980) observation that there is a stage in Voice Onset Time (VOT) acquisition in which children produce target voiceless stops with significantly longer VOTs than target voiced stops but both within the short lag region; a listener, however, may not perceive the difference between these two stops. Other authors have reported covert contrast in the acquisition of place of articulation (PoA) such as the alveolar-velar stop distinction (Forrest et al., 1990; McAllister Byun et al., 2016), frontal misarticulations of /s/ ([s] vs. [θ], Schellinger et al., 2017) and in the alveolar-alveopalatal fricative distinction (Li et al., 2009), which we will discuss in greater detail.

Spectral moments' analysis of the alveolar-alveopalatal distinction

Li et al. (2009) used spectral moments' analyses to examine covert contrast in the acquisition of /s/ and /ʃ/ in English- and Japanese-speaking children. A spectral moments' analysis computes mathematical moments from the power spectrum. Four spectral moments are generally considered; however, in this study, we focus on the first spectral moment (also referred to as the spectral mean or centroid) which is one of the most useful for distinguishing alveolar and alveopalatal PoA. It calculates the average energy concentration, which is related to the location of constriction in the oral cavity. The point of constriction for /s, z/ is more anterior than for /ʃ, ʒ/ resulting in a shorter frontal cavity and higher mean energy.

Li et al. (2009) analyzed /s/ and /ʃ/ produced in word-initial position in two- and three-year-old English- and Japanese-speaking children. Given the fact that the children produced many fricatives with errors, they separated out children who produced a contrast between alveolar and alveopalatal fricatives and those who did not on the basis of phonetic transcription. They found a significant difference between /s/ and /ʃ/ (or between /s/ and /ɕ/ in the case of Japanese) for the first spectral moment in those children who were transcribed as producing a contrast; however, the distinction between the two fricatives was smaller than observed in adults. In those children who were transcribed as making errors, there was a covert contrast in a small number of children (n=4), although not necessarily for the first spectral moment (a covert contrast was evident for other spectral moments or for the onset of the second formant). In the remainder of the children (n=12),

there was no evidence of a distinction between /s/ and /ʃ/ in the transcription or spectral moments' analyses. Li et al. (2009) tested very young children (i.e., two- to three years) and it might be the case that covert contrast is more pronounced in older children. In addition, Li et al. (200) employed the criteria of 75% mastery to distinguish children who did or did not have a contrast in their phonetic transcription. However, this distinction may not be fine enough to classify children who don't have a distinction since there may be differences in spectral realization at low and middle levels of accuracy.

In this study, we examine whether French-speaking children, exhibit covert contrast between /s/ and /ʃ/ in their spectral moments' realization. We test children between two and six years of age to determine whether age influences the presence of covert contrast. Furthermore, we examine the presence of covert contrast in children who vary in their alveopalatal fricative accuracy. The finer categorization of alveopalatal fricative mastery may provide additional information on when children evidence covert contrast in their speech.

Methodology

Participants

Participants included 80 French-speaking monolingual and bilingual children, aged 2;6 to 6;10. Information on the number of participants across age is provided in Table 1. Based on questionnaire information, 20 of the 39 bilingual children were dominant and 19 were not dominant in French. The children had all received exposure to French before the age of three years. The languages spoken by the bilinguals included English, German, Swedish, Italian, and Spanish.

Table 1. Numbers of participants across age

Age	Mean age	n	Mon[a]	Bi	Girls	Boys
2	2;6	17	9	8	7	10
3	3;4	15	8	7	5	10
4	4;6	16	8	8	10	8
5	5;3	16	8	8	10	8
6	6;4	16	8	8	7	9

[a] Mon = monolinguals; Bi = bilinguals

Table 2. Examples of word stimuli containing /s/ and /ʃ/

Sound	Word-initial	Word-medial	Word-final
s	cinq	chaussure	brosse
	six	dessin	glace
ʃ	chaise	cochon	vache

chat	échelle	bouche

Stimuli

The stimuli for the children included words ranging from one to three syllables with target /s/ and /ʃ/ situated in word-initial, -medial, or -final positions. The majority of words were familiar to children as young as two- to three years. Examples of the word stimuli are provided in Table 2.

Procedure

Children took part in an object or word naming task of approximately 20 to 30 minutes (see Kehoe & Girardier, 2020 for further details). The two-year-olds were tested in the speech laboratory at the University of Geneva and the older children were tested in a quiet room in the children's kindergarten or school. Children's productions were recorded with a portable digital tape recorder.

Data Analyses

Using Phon, a software program designed for the analysis of phonological data (Rose & MacWhinney, 2014), each child's WAV file was segmented, and stimulus words were identified and transcribed. French-speaking graduate students, who had experience in phonetic transcription, performed the analyses.

Acoustic analyses were conducted in Praat (Boersma & Weenink, 2016). We used the time waveform, spectrogram, and amplitude contour to aid in the segmentation of fricatives. Once the fricative segment was identified, we ran a Praat script that extracted six spectra across the length of the fricative, averaged these spectra, and computed spectral moments based on this averaged spectrum. A large number of productions (approximately 300) were excluded because they were characterized by noise overlay, low or high volume, or were tokens that were difficult to segment.

Data-coding and Statistical Analyses

The analyses were performed using R statistical software (R Development Core Team, 2020) and the lme4 package for mixed models. The dependent variable was the spectral mean or centroid. Independent variables included age (in months), gender (male, female), bilingual status (mon, bi), sound (s or ʃ), and /ʃ/ accuracy. We used a continuous variable of percent accuracy based on the phonetic transcription of words included in the final database. In addition, we examined the interaction between sound and age to determine whether the centroid contrast between /s/ and /ʃ/ increased with age, and the interaction between /ʃ/ accuracy and sound to determine whether the centroid contrast was dependent upon the accuracy of /ʃ/. Random factors included participant and word token.

Results

Percent correct accuracy of /s/ and /ʃ/ across age range is presented in Figure 1. Percent correct accuracy was high for /s/ at all age ranges whereas accuracy for /ʃ/ was low at the youngest age

and grew steadily until it approached 100% at 6 years. Standard deviations were large for /ʃ/ between two to five years.
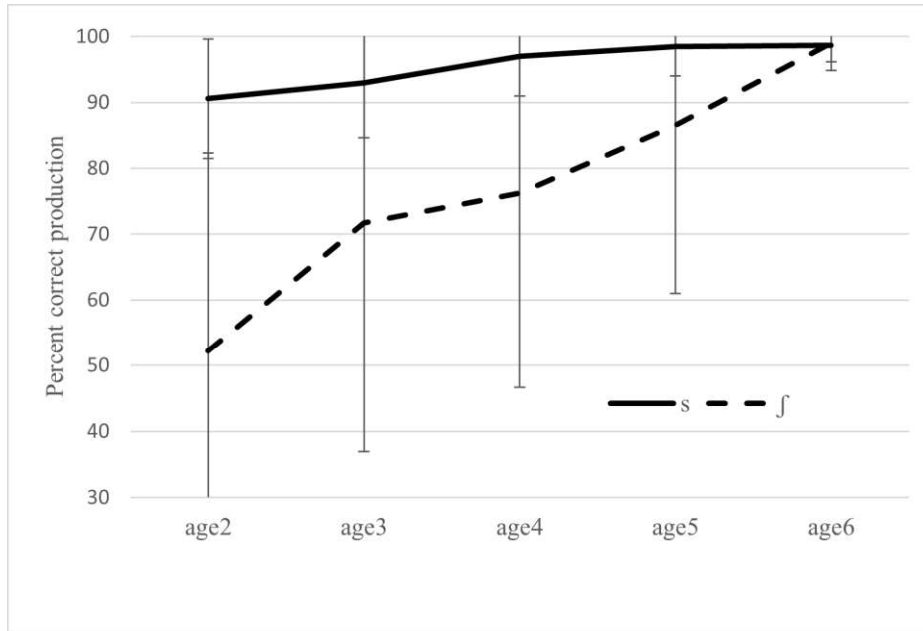

Figure 1. Percent correct production of /s/ and /ʃ/ across age. Error bars indicate standard deviations

Table 3 presents the best-fitting model for factors that influence centroid values. Age, gender, and sound were significant. Centroid values decreased with age, were higher in girls than boys, and were higher in /s/ versus /ʃ/. There was no influence of monolingual status on centroid realization. In addition, there was a significant interaction between age and sound: the difference between centroid values for /s/ and /ʃ/ increased with age. There was also a significant interaction between /ʃ/ accuracy and sound. As shown in Figure 2, the spectral mean contrast between /s/ and /ʃ/ increased with increasing /ʃ/ accuracy.
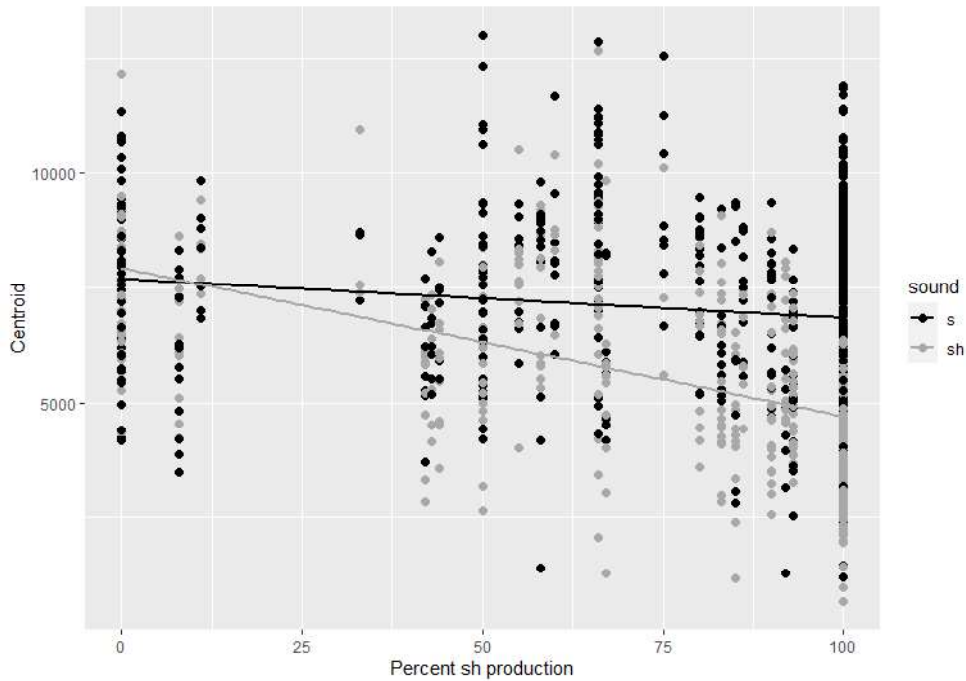
Figure 2. Centroid values for "s" and "sh" according to percent /ʃ/ production (based on phonetic transcription)

Table 3. Best fitting model for factors influencing centroid values of /s/ and /ʃ/

Fixed Effects	Estimate	Std. Error	t value		Pr(>|t|)
(Intercept)	9274.883	418.215	22.177		< .001 ***
age	-32.467	6.857	-4.735		< .001 ***
gender	-509.540	197.423	-2.581	0.01 *	
MonBi	-39.943	195.244	-0.205	0.84	
sound	940.912	314.643	2.990		0.003 **
/ʃ/ accuracy	-3.316	3.627	-0.914	0.36	
age:sound	-11.894	5.007	-2.375	0.02 *	
/ʃ/ accuracy	-19.730	2.757	-7.157		< .001 ***

To obtain additional information on the nature of the contrast, we ran a second model categorizing children into low (0-20%), middle (33-67%), and high (75-100%) [ʃ] producers to examine whether a spectral mean contrast was present for each group. The centroid means for each group of children are shown in Figures 3 to 5. Tukey multiple comparisons indicated that low /ʃ/ producers (n=8) did not make a significant centroid contrast between /s/ and /ʃ/ (t=-.25, p=.99).

The mean values for /s/ and /ʃ/ were 7338.75 (sd=1882.29) and 7361.89 (1725.20). However, the mid (n=15; t=3.5, p=.007), and high (n=57: t=12.85, p<.001) producers did. The mean values for /s/ and /ʃ/ were 7642.77 (2022.59) and 6643.66 (1963.44) in the mid group and 6800.26 (1853.45) and 4728.17 (1411.42) in the high group. The difference between /s/ and /ʃ/ increased with age in the high but not in the mid producers.

Discussion and Conclusion

This study conducted a spectral moments' analysis of /s/ and /ʃ/ in French-speaking children, aged 2;6 to 6;10. The children varied in their mastery of [ʃ] based on phonetic transcription, often neutralizing the distinction between /s/ and /ʃ/. We aimed to determine whether children who did not display any perceptible differences between /s/ and /ʃ/ were nevertheless producing an acoustic contrast in their spectral moments. Our findings indicated that children who were rarely perceived as producing [ʃ] (i.e., accuracy less than 20%) did not distinguish /s/ and /ʃ/ on the basis of spectral moments. These children ranged in age from two to five years. Thus, regardless of age, covert contrast was not evident in their speech. Children who were starting to produce [ʃ] (i.e., accuracy between 33 and 66%) distinguished /s/ and /ʃ/ on the basis of spectral moments. These children were also aged between two to five years, and the magnitude of the centroid contrast was similar across age. Finally, children who had mastery of [ʃ] (i.e., accuracy greater than 75%), exhibited a spectral mean contrast that increased across age. They ranged in age from two to six years. Thus, the findings suggest that a centroid contrast starts to develop in children once they are perceived as producing a certain number of alveopalatal fricatives (e.g., approximately 30%).
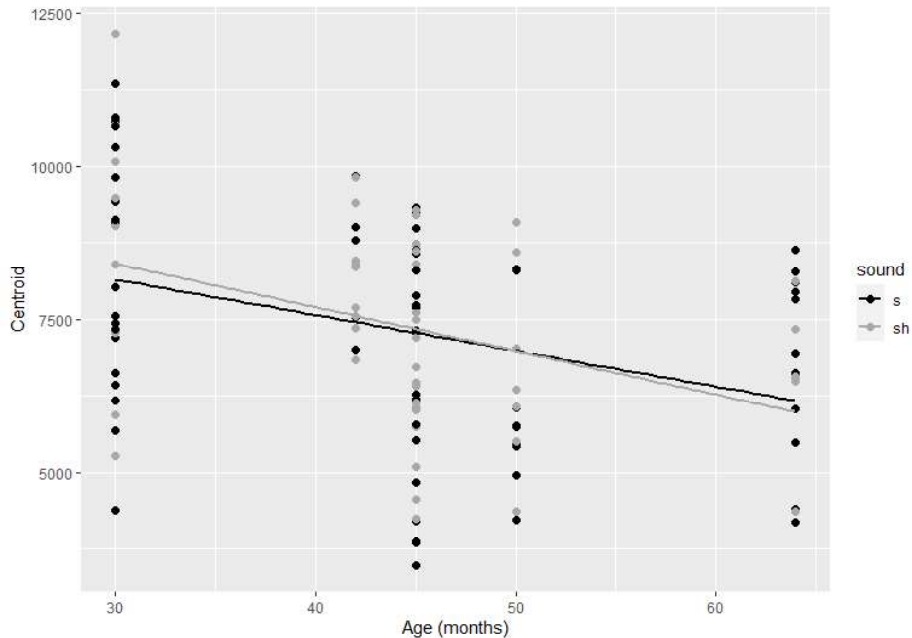

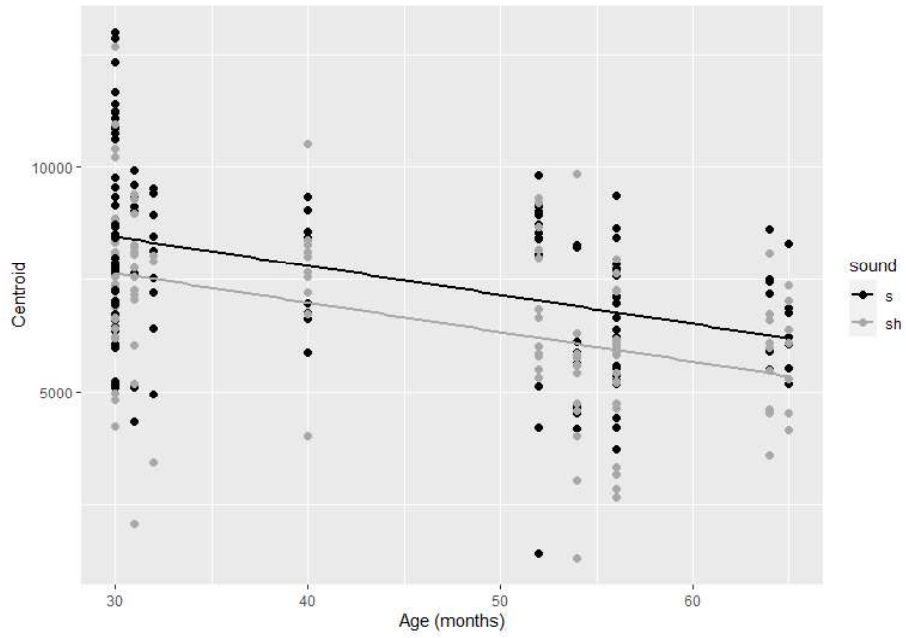**Figure 3. Centroid values for "s" and "sh" in the low [ʃ] producers (n=8) across age**

Figure 4. Centroid values for "s" and "sh" in the [ʃ] producers (n=15) across age
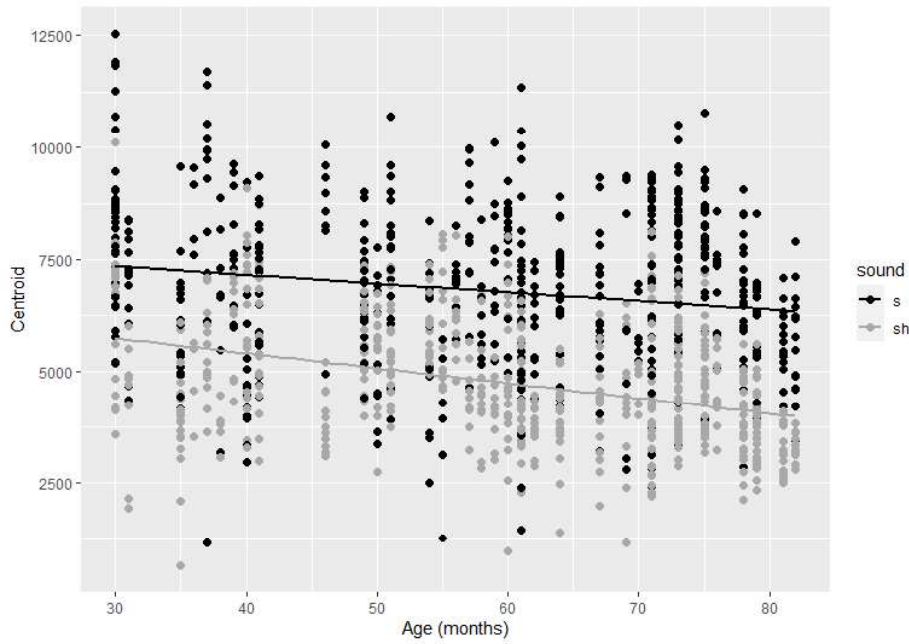

Figure 5. Centroid values for "s" and "sh" in the high [ʃ] producers (n=57) across age

This study has several limitations including the fact that the number of children who were low and middle [ʃ] producers was small in the current sample. The sample included predominantly children who were high [ʃ] producers. The low percent accuracy for [ʃ] evident in Figure 1 stems from the fact that, at ages two to five years, there were children who varied between 0 and 100% production. A sample with a greater number of low [ʃ] producers may have revealed stronger evidence of covert contrast. Furthermore, we concentrated on voiceless fricatives, and the inclusion of voiced fricatives [z, ʒ], which are articulatorily more challenging, may have provided greater evidence of covert contrast, as might have the analysis of other spectral moments (e.g., skewness) apart from the centroid. Nevertheless, our findings revealed that many children display a reliable acoustic contrast between [s] and [ʃ] even though their transcription data suggest that the contrast has not yet been mastered.

References

Aicart-De Falco, S. & Vion, M. (1987). La mise en place du système phonologique du français chez les enfants entre 3 et 6 ans: Une étude de la production. *Cahiers de Psychologie Cognitive*, *7*, 247-266.

Boersma, P., & Weenink, D. (2016). PRAAT, a system for doing phonetics by computer (version 6.0. 20) [software package]. Institute of Phonetic Sciences, University of Amsterdam, The Netherlands.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics and Phonetics*, *4*(4), 327-340. https://doi.org/10.3109/02699209008985495

Kehoe, M., & Girardier, C. (2020). What factors influence phonological production in French-speaking bilingual children, aged three to six years? *Journal of Child Language*, *47*(5), 945-981. https://doi.org/10.1017/s0305000919000874

Kehoe, M., & Havy, M. (2019). Bilingual phonological acquisition: The influence of language-internal, language-external, and lexical factors. *Journal of Child Language, 46*(2), 292-333. https//:doi.org/10.1017/S0305000918000478

Li, F., Edwards, J., & Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*, *37*(1), 111-124. https://doi.org/10.1016/j.wocn.2008.10.001.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*(1), 41-74. https://doi.org/10.1017/s0305000900007029

McAllister Byun, T., Buchwald, A., & Mizoguchi, A. (2016). Covert contrast in velar fronting: An acoustic and ultrasound study. *Clinical Linguistics and Phonetics*, *30*, 249-276. https://doi.org/10.3109/02699206.2015.1056884

MacLeod, A. A. N., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in Québécois French: A cross-sectional study of pre-school aged children. *International Journal of Speech-Language Pathology, 13*, 93-109. https://doi.org/10.3109/17549507.2011.487543

R Development Core Team, (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna.

Rose Y., & MacWhinney B. (2014). The PhonBank initiative. In J. Durand J, U. Gut, & G. Kristoffersen (Eds.). *The Oxford handbook of corpus phonology* (pp. 380-401). Oxford University Press.

Schellinger, S. Munson, B. & Edwards, J. (2017). Gradient perception of children's productions of /s/ and /θ/: *A comparative study of rating methods, Clinical Linguistics and Phonetics, 31*, 80-103. https://doi.org/10.1080/02699206.2016.1205665