PRODUCTION OF POLISH SIBILANTS IN THE PROCESS OF LANGUAGE ACQUISITION

Marzena Żygis[1], Marek Jaskuła[2], Daniel Pape[3] & Laura Koenig[1,4,5]

[1]Leibniz-Centre General Linguistics, Berlin, Germany; [2]Westpomeranian University of Technology, Szczecin, Poland; [3]McMaster University, Hamilton, Canada; [4]Adelphi University, Garden City, New York; [5]Haskins Laboratories; New Haven CT, USA

zygis@leibniz-zas.de, jaskula@Marek.Jaskula@zut.edu.pl, paped@mcmaster.ca, koenig@haskins.yale.edu

ABSTRACT

This paper reports acoustic characteristics of Polish sibilants acquired in the process of language learning. We tested the production of three phonemic sibilants /s, ʂ, ɕ/ produced by 81 Polish children ages 35 to 106 months.

Our results based on an acoustic analysis complemented by a perceptual categorization test by adults reveal that the alveolo-palatal /ɕ/ is the first sound which becomes separated from the other sibilants in terms of F2 of the following vowel and centre of gravity. The next sounds acquired are /s/ and /ʂ/. In the perceptual test most errors were found for the retroflex /ʂ/ confirming its late acquisition in comparison to other sibilants.

The order of acquisition in Polish partly differs from that of Putonghua (/ɕ/ > /ʂ/ > /s/) which is explained by different frequency of the phonemes /ʂ/ and /s/ in both languages.

Keywords: sibilants, acoustics, Polish, acquisition.

1. INTRODUCTION

While some research has been conducted on acquistion of simple sibilant contrasts, very little is known about how children acquire complex sibilant systems. Previous experimental research mainly focused on English /s, ʃ/ ([3], [7], [8], [9], [13], [17], [19], [24]) and Japanese /s, ɕ/ ([7], [8], [9]) with the exception of the study by Li [7] and Li and Munson [10] who examined production and perception of Putonghua (Mandarin) three-way contrast /s, ʂ, ɕ/. They could show that the order of sibilant acquisition was /ɕ/ > /ʂ/ > /s/ while, as stated in [8], it was /s/ > /ʃ/ in English and /ʃ/ > /s/ in Japanese.

Li [7] also showed that the acquisition of a complex sibilant inventory varies considerably from that of a simple one. English and Japanese children significantly differed in the acoustic dimensions that contrasted the fricative categories. While differences in the first spectral moment (centre of gravity) were decisive for the acquisition of the simple contrasts, it was not only the first spectral moment but also the second moment and the frequency of the second formant at the following vowel onset which were of particular relevance for the acquisition of the complex contrast.

However, it has remained unexplored whether there are similarities in the acquisition of complex contrasts among languages or whether the acquisition process is language-specific.

To address this question we probed the production of the complex three-way phonemic contrast /s, ʂ, ɕ/ by 81 Polish children as function of their age. According to our knowledge this is the first acoustic study on Polish sibilant acquisition on a large scale.

2. EXPERIMENT

2.1. Experimental design and set up

The experiment consisted of three parts. In the first, the children were asked to name pictures displayed on a screen. The second and third parts of this experiment were testing the perception of sibilants. The present paper focuses on the production results obtained in the first part of the experiment. In this part, preceded by a short training section, children were shown pictures and they were asked to name the object. Three of the objects contained the sibilants /s, ʂ, ɕ/ in the word-medial position which created a triplet, i.e. /kasa/, "cash point" and /kaɕa/ "Katie, women's nickname" and /kaʂa/ "groats". Other words included the same sibilants in the word-initial position: /sanki/ "sleigh", /ɕatka/ "net", /ʂafa/ "wordrobe". Finally, there were three distractors: /rɨba/ "fish", /swɔnkɔ/ "sun", and /ʐaba/ "frog".

The pictures were shown three times, in a randomized order for each participant. Thus, we obtained 27 word recordings for each child (9 words x 3 repetitions).

All words were bisyllabic and accented on the first syllable. They were also of similar frequency and the children did not encounter difficulties in naming the pictures.

It is worth mentioning that for the purpose of this (and following) experiments we built an Ubuntu Mate based tool which we called *Linguistino* with the help of which we are able to use our own design, record children' voices and adjust the answer box (colours and placements of the buttons) in the perceptual part of our experiment.

2.2. Informants

The recordings were made with 81 preschool and school-aged children (39 female), all monolingual native speakers of Standard Polish living in Szczecin (Northern Poland). Their ages varied from 35 to 106 months.

The parents of the participants signed a consent after being informed about the experiment in detail. As a compensation for their effort each child could choose a souvenir.

2.3. Procedure

Our acoustic analysis exclusively concentrates on the sibilants appearing in the word-medial position in the triplet: /kasa/, “cash point”, /kaɕa/ “Katie” and /kaʂa/ “groats”.

We computed acoustic multitaper spectra with a 512 point Hamming window at the acoustic midpoint of each sibilant (see [26] for the advantages of multitaper over other analysis methods). The midpoint was defined as the acoustic landmark between the onset and offset of the (voiceless) sibilant frication noise. The power spectral density (PSD) was estimated via the Thomson multitaper method (linear combination with unity weights of individual spectral estimates and the default Fast Fourier Transform (FFT) length) available in the MathWorks Signal Processing Toolbox Version 6.2 [14]. The highest spectral peak frequency was computed for the full frequency range from 0 Hz to 22050 Hz.

Furthermore, for the full frequency range we computed the four spectral parameters based on the formulae given in Praat version 5.2 ([1]): centre of gravity (COG), standard deviation of the spectrum, skewness, and kurtosis of the spectrum.

Next, formant frequencies F1, F2, F3 were measured at the end of the preceding and at the beginning of the following vowel using PRAAT’s automatic formant extraction algorithm („To Formant... burg“) with standard parameter settings (max. formants: 5500Hz; max. number of formants: 5; window length: 25 ms, Pre-emphasis from 50Hz), i.e. the first three extracted spectral maxima of the described Praat algorithm were considered to be the formants F1-F3. The data were not manually checked.

For each CVCV item, we calculated the acoustic durations of the sibilant and the previous and following vowels.

The sounds were also perceptually categorized by two native speakers of Polish. They listened to words produced by children and judged whether they heard /s/, /ʂ/ and /ɕ/ or another sound. If they heard the sound intended by the child the answer was classified as correct; otherwise as incorrect.

2.4. Statistics

All statistical analyses were conducted in the R Studio software (version 1.1.453, [20]).

A multinomial analysis testing the influence of the following parameters on the variable SOUND with three sublevels [/s/, /ʂ/, /ɕ/] was conducted by including the following parameters: (a) F1, F2 and F3 at the end of the preceding vowel, (b) F1, F2 and F3 at the beginning of the following vowel, (c) duration of the sibilant, preceding and following vowel, (d) frequency of the highest peak, cog, standard deviation, skewness, kurtosis, (e) speaker’s sex [male, female], and (f) age. We also included interactions of age with parameters (a)-(e). Prior to the analysis all predictors were centered to reduce multicollinearity.

3. RESULTS

First of all, we observe a huge inter-speaker variation with respect to all sibilants. Figures 1, 2 and 3 present multitaper spectra of /s/, /ɕ/ and /ʂ/ from all informants excerpted at the midpoint of frication. The spectra for individual productions are shown by green lines and the overlaid mean spectrum is represented by the black line.

Figure 1: Spectra obtained for /s/ productions

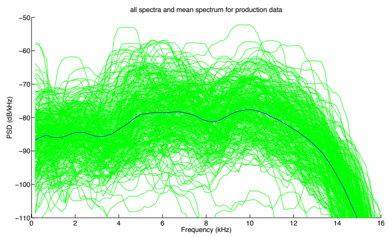


Figure 2: Spectra obtained for /ɕ/ productions

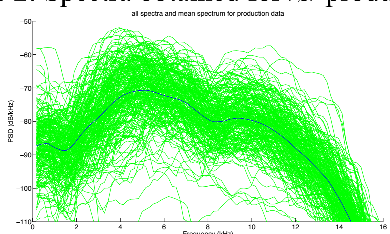
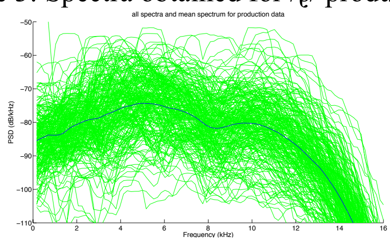


Figure 3: Spectra obtained for /ʂ/ productions



3.1. Acoustic analysis

In order to examine how the sibilants develop over time we divided the results according to three age groups, as presented in Table 1:

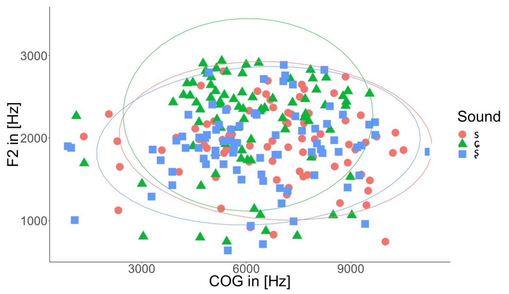
Table 1: Children groups

	Age (months)	Number of children	Number of observations
Group 1	35-50	27	239
Group 2	51-70	37	338
Group 3	71-106	16	144

We first provide results on measures which, based on previous literature ([18], [25]), appear to be most decisive in the classification of sibilants: F2 measured at the beginning of the following vowel and centre of gravity (COG) of the spectrum.

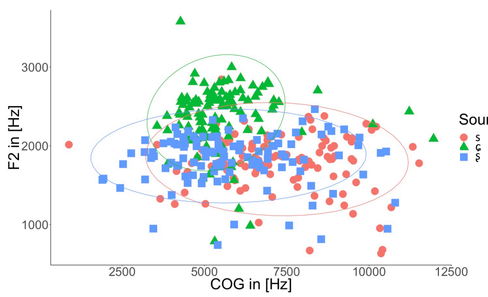
Figure 4 presents results for the youngest group of children. It appears that the sibilants are almost randomly spread. It also seems, however, that the alveolopalatal [ɕ] is least spread and as such may have started to separate from other sounds.

Figure 4: COG and F2 frequency of the sibilants for the first group (/s/=circle, /ɕ/=triangle, /ʂ/=square)



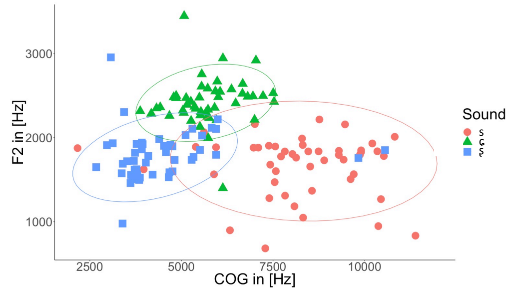
In the second group, where children have already finished the fourth year of life, it turns out that the alveolo-palatal is indeed the least spread sound and its separation from other sounds continues. This is presented in Figure 5.

Figure 5: COG and F2 frequency of the sibilants for the second group



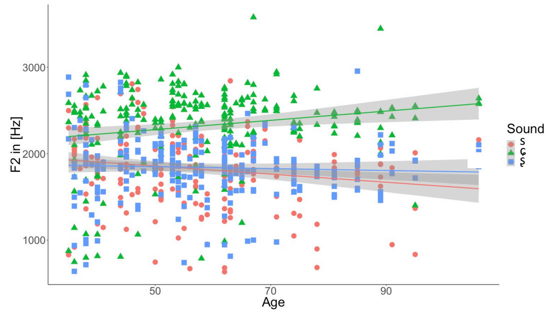
Finally, as illustrated in Figure 6, sibilants produced by the eldest children are well separated. A few sibilant productions still overlap, but it should considered that only two acoustic parameters, i.e. COG and F2 frequency, have been taken into account so far (see section 3.2 for other parameters).

Figure 6: COG and F2 frequency of the sibilants for the third group



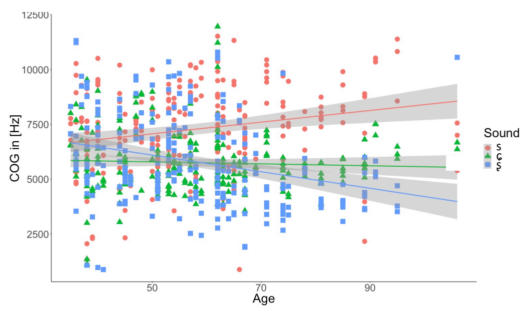
If we consider F2 independently of COG it turns out that this formant separates the alveolo-palatal right at the beginning of acquisition process; see Figure 7.

Figure 7: F2 of the vowel following /s, ʂ, ɕ/ as a function of age.



By contrast, COG results for individual sibilants show that this spectral parameter does not contribute to the differentiation of the sounds at the very beginning because the parameter overlaps for all sibilants. However, it starts to contribute to the separation of /s/ shortly before 50 months. This also suggests that /s/ is acquired before /ʂ/; see Figure 8.

Figure 8: COG of /s, ʂ, ɕ/ spectra as a function of age.



3.2. Multinomial analysis

Table 2 presents the results of a multinomial analysis revealing a significant interaction of a given parameter with age with respect to sound pairs.

Table 2: Results of a multinomial analysis

	coeff.	z value	p value
a) /s/ vs. /ʂ/:			
F2 of V1*Age	-0.345	-2.28	p<.05
F1 of V2*Age	-0.257	-1.99	p<.05
F3 of V2*Age	0.493	3.17	p<.01
COG*Age	0.523	2.01	p<.05
STD*Age	0.399	2.53	p<.05
kurtosis*Age	0.430	2.33	p<.05
b) /s/ vs. /ɕ/:			
F1 of V2*Age	0.362	2.30	p<.05
F2 of V2*Age	-1.793	-7.42	p<.001
F3 of V2*Age	0.896	4.29	p<.01
STD*Age	0.840	3.77	p<.001
Kurtosis*Age	0.672	2.84	p<.01
c) /ʂ/ vs. /ɕ/:			
F1 of V2*Age	0.57	3.93	p<.001
F2 of V2*Age	-1.55	-6.85	p<.001
F3 of V2*Age	0.56	2.81	p<.01
STD*Age	0.42	1.97	p<.05

The analysis shows that several parameters play a significant role in the process of sibilant acquisition in Polish. In particular, it turns out that not only F2 but also F1 and F3 at the following vowel onset differentiate the alveolo-palatal /ɕ/ from /s/ and /ɕ/ in acquisition. For the /ʂ/ vs. /s/ contrast, F2 of the following vowel turns out to be not significant, but instead spectral properties such as COG, STD and kurtosis (as well as other formants) reflect acquisition of the contrast.

3.3. Perceptual analysis

Table 3 presents the results of the perceptual test. The inter-rater agreement calculated as Cohen's kappa was 0.71.

Table 3: Results of the perceptual test

	/s/	/ʂ/	/ɕ/
Annotator 1:			
correct	163 (67.6%)	102 (42.5%)	204 (85%)
3.1. Annotator 2:			
correct	185 (76.8%)	123 (51.3%)	186 (77.5%)

The results show that the least correct answers were given when the child intended to pronounce the retroflex /ʂ/ suggesting that it is the most difficult sound for children to produce correctly. The highest number of correct answers was assigned to /ɕ/ suggesting that it was the easiest sound. However, for the second annotator the difference between /ɕ/ and /s/ was minimal.

4. DISCUSSION AND CONCLUSIONS

Our results reveal huge inter-speaker variation in the acquisition of all sibilants which is in line with studies on sibilant acquisition in other languages ([7], [9]).

However, based on our acoustic data complemented by perceptual results, we were able to detect a clear acquisition pattern: /ɕ/ > /s/ > /ʂ/. This order is in line with previous studies, based either on impressionistic perception or a small number of children, see also [5], [12] and [11] for an overview of different studies.

Comparing the order of acquisition of Polish sibilants to that of Mandarin it turns our that /ɕ/ is the first acquired sibilant in both languages. However, the second sibilant in Mandarin is the retroflex /ʂ/ and the latest sibilant is /s/.

Following [10] and [7] we hypothesize that the early acquisition of the alveolo-palatal is due to motor control mechanisms: /ɕ/ is produced with the tongue dorsum and control over this articulator should precede that for fine movements of the apex or tongue blade.

Regarding the different acquisition order of the other sibilants, we put forward a hypothesis that it reflects the frequency of occurrence of the sibilants in those languages. As shown by [10], the frequency order of all sibilants in Putonghua is /ʂ/ > /ɕ/ > /s/), i.e. /ʂ/ occurs more frequently than /s/ does. On the contrary, in Polish, the frequency order of all sibilants is /s/ > / ɕ/ > /ʂ/, i.e. /s/ occurs more often than /ʂ/; see [6]. In addition, although quantitative studies are still missing, it has been reported that /ɕ/ is commonly used in motherese in Polish [4].

Finally, the early acquisition of /ɕ/ can be explained by the fact that transitions of vowels play a crucial role in the production (and perception) of fricatives because younger children relay more on dynamic than static information in the process of phonemic categorization ([15], [16]). This might also explain why the process of separation of /s/ and /ʂ/ takes place later: the transitions are not as salient as in the case of /ɕ/, see also [2], [21], [22], [23]. Further perceptual studies are needed to validate this hypothesis.

Acknowledgements
Marzena Żygis was supported by the BMBF (grant number 01UG1411) and the Leibniz Association. We also thank Ania Gogolin for drawing the pictures for the experiment, and all our informants for taking part in this study.

5. REFERENCES

[1] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org/

[2] Bukmaier, V., Harrington, J. 2016. The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *Journal of the International Phonetic Association* 46, 311-329.

[3] Cristià, A. 2010. Phonetic enhancement of sibilants in infant-directed speech. *The Journal of the Acoustical Society of America* 128, 424-434.

[4] Czaplicki B., Żygis, M., Pape D. Jesus, L. 2016. Acoustic evidence of new sibilants in the pronunciation of young Polish women. *Poznań Studies in Contemporary Linguisitics* 52, 1-42.

[5] Kaczmarek, L. 1953. *Kształtowanie sie mowy dziecka*. [Development of child speech] Poznań: Poznańskie Towarzystwo Przyjaciół Nauk.

[6] Kłosowski, P. 2017. Statistical analysis of orthographic and phonemic language corpus for word-based and phoneme-based Polish language modelling. *EURASIP Journal of Audio, Speech, and Music Processing 5.* https://doi.org/10.1186/s13636-017-0102-8

[7] Li, F. 2008. *The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese*. PhD dissertation, The Ohio State University.

[8] Li, F. 2012. Language-specific developmental differences in speech production : A cross-language acoustic study. *Child Development* 83, 1303-1315.

[9] Li, F., Edwards, J., Beckman, M. 2009. Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics* 37, 111–124.

[10] Li, F., Munson, B. 2016. The development of voiceless sibilant fricatives in Putonghua-speaking children. *Journal of Speech, Language, and Hearing Research* 59, 699-712.

[11] Łobacz, P. 1996. *Polska fonologia dziecięca* [Polish child phonology]. Warszawa: Energeia.

[12] Łukaszewicz, B. 2000. *Phonological analysis of Reduction and Assimilation Processes in the Acquistion of English and Polish*. PhD dissertation, University of Warsaw.

[13] Mann, V., Sharlin, H.M., Dorman, M. 1985. Children's perception of sibilants: the relation between articulation and perceptual development. *Journal of Experimental Child Psychology* 39, 252-264.

[14] MathWorks. 2007. Signal Processing Toolbox 6 User's Guide. Natick: MathWorks, 470–475.

[15] Nittrouer, S. 2002. Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustical Society of America* 112, 711–719.

[16] Nittrouer, S., Studdert-Kennedy, M. 1987. The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research* 30, 319–32.

[17] Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., Tiede, M. 2018. How do children organize their speech in the first years of life ? Insight from ultrasound imaging. *Journal of Speech, Language and Hearing Research* 61, 1355-1368.

[18] Nowak, P. M. 2006. The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Speech and Hearing Research* 34, 139–152.

[19] Reidy, P., Beckman, M., Edwards, J., Munson, B. 2017. A data-driven approach for perceptually validated acoustic features for children's sibilant fricative productions. *Proceedings of Interspeech*, Stockholm.

[20] RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Version 1.1.453.

[21] Vihman, M. 2015. Perception and production in phonological development. In: MacWhinney, B., O'Grady, W. (eds.) *The Handbook of Language Emergence*. Oxford, UK: Wiley Blackwell. 437-457.

[22] Walley, A. 2008. Speech perception in childhood. In: Pisoni, D.B., Remez R. (eds.) *The Handbook of Speech Perception*. Wiley Online Library.

[23] Werker, J.F., Tees, R.C. 1999. Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology* 50, 509-535.

[24] Zharkova, N., Hardcastle, W.J., Gibbon, F.E. 2018. The dynamics of voiceless sibilant fricative production in children between 7 and 13 years old. An ultrasound and acoustic study. *Journal of the Acoustical Society of America* 144, 1454–1466.

[25] Żygis, M., Padgett, J. 2010. A perceptual study of Polish sibilants, and its implications for historical sound change. *Journal of Phonetics* 38, 207-226.

[26] Żygis, M., Pape, D., Jesus L. 2012. (Non)retroflex Slavic affricates and their motivation. Evidence from Czech and Polish. *Journal of the International Phonetic Association* 42, 281-329.