Breathiness in speech directed to 4-month-old infants

Anna Kohári[1], Uwe D. Reichel[1], Alexandra Markó[2,3] & Katalin Mády[1]
[1]*MTA Research Institute for Linguistics,* [2]*Eötvös Loránd University,* [3]*MTA-ELTE "Lendület" Lingual Articulation Research Group*

Introduction. Mothers tend to speak differently to infants than to adults. This register is referred to as motherese or infant-directed speech (IDS), whereas the one used talking to adults is called adult-directed speech (ADS) [1]. Higher fundamental frequency and slower speech rate are the most typical characteristics of IDS compared to ADS [1, 2]. There is rather sparse data available on the voice quality features of IDS, and it has not been in the focus of previous research. Meanwhile, in other areas of speech entrainment studies voice quality is gaining increasing attention. Its several measurable acoustic parameters make voice quality a useful indicator of speech entrainment [3] and facilitate the identification of positive emotion expressions [4], which are fundamental features of IDS. A recent study has found that on average vowels were produced with more breathy voice quality in Japanese speech directed to 20-month-old infants than in ADS [5]. However, it remains unknown whether the occurrence of breathiness is a universal feature of motherese and whether it already appears in the first year of the infants' lives or at a later stage. The present study aims to investigate the voice quality characteristics of Hungarian speech directed to 4-month-old infants.

Methods. Recordings of 20 Hungarian native speakers were investigated. Mothers were given a book containing pictures and sentences. They had to create a story using their own words and also including utterances written in the pictures. First, they had to present the story to the experimenter (ADS), afterwards they repeated the same process to their 4-month-old infants (IDS). We investigated the vowels /ɒ/, /a:/, /ɛ/, /i/, /o/ and /u/, whose occurrences in the 9 fixed sentences were frequent enough for statistical analyses. These vowels could be found in at least 5 tokens, which included both stressed and unstressed realizations of all vowels in various word positions. We used MAUS [6] for segmenting the utterance and vowel boundaries automatically and corrected them manually in the Praat 6.0 software. To investigate voice quality, we measured cepstral peak prominence (CPP), the difference between the first two spectral harmonic magnitudes (H1−H2) and its formant-corrected counterpart (H1*−H2*) and harmonics-to-noise ratios (HNRs) in various frequency bands (0–500Hz, 0–1500Hz, 0–2500Hz) using VoiceSauce [7]. These metrics are related to the perception and production of breathiness. The lower CPP values indicate that vowels are less modal, whereas the higher values of H1*−H2* and HNR imply breathy voice [8]. As the surrounding consonants can influence these parameters [8], the investigated segments were split into 3 sections of equal length, the mid third of each vowel was analyzed. Statistical analysis was carried out in R. We constructed separate mixed effect models for each vowel with the fixed effect register (IDS vs. ADS) and random slopes by subjects. (We also built mixed effect models including random intercepts by items yielding practically the same results.)

Results. We found that H1*−H2* values were significantly higher in IDS than in ADS for almost all vowels ($p < 0.05$) except for /i/ and /u/ where IDS did not differ significantly from ADS. CPP values were lower in IDS ($p < 0.05$) than ADS but here, too, /i/ and /u/ did not show differences between the two registers (Figure 1). HNR metrics appeared to be unaffected by the registers. All analyses were repeated excluding the vowels of utterance-final syllables, which led to similar results.

Conclusions. IDS was found to be characterized by breathy voice compared to ADS based on CPP and H1*−H2* values. It cannot be ruled out that the lack of significant differences between the two registers in the case of /i/ and /u/ can be explained by the effect of the

surrounding consonants, however, vowel height can also influence the values of voice quality metrics. High vowels tend to have higher H1*−H2* [9], which may have blurred voice quality differences between the two registers.

To summarize, IDS was found to be breathier than ADS in Hungarian, a phenomenon that appears to be a general feature across various languages. In IDS, breathy voice was markedly detectable in the early months of the infants' age. Breathy voice occurs more often in speech conveying a happy and calm emotional state of the speaker [4, 10], therefore breathiness in IDS can be interpreted as the expression of positive emotions.

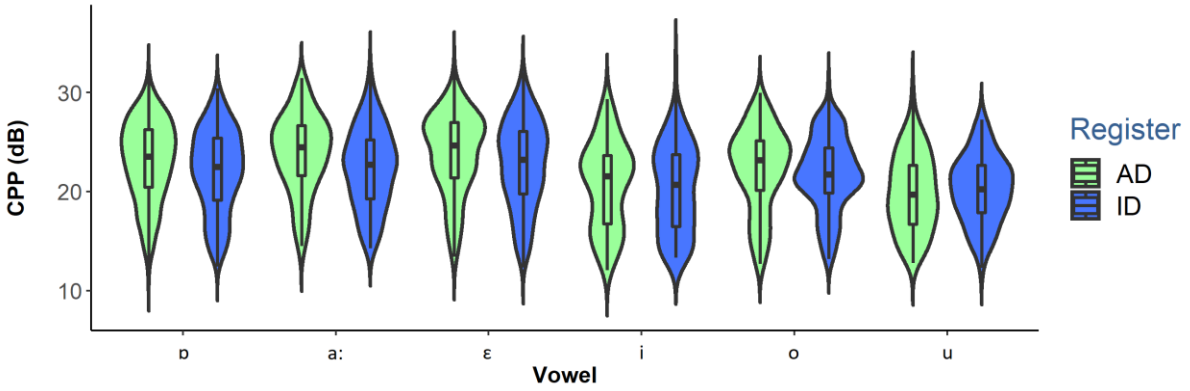

Figure 1. *CPP values for each investigated vowel in the two registers*.

This work was supported by the Hungarian National Research, Development and Innovation Office (NKFIH) under grant 'Neurocognitive predictors of early language development' (no. NKFI-115385) and 'Longitudinal study of infant-directed speech using multimodal methods' (no. NKFI-134775).

References

[1] Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M-C., & Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and cognition?(A systematic review). PloS one, 8(10), e78103.

[2] Mady, K., D Reichel, U., Szalontai, A., Kohari, A., & Deme, A. (2018). Prosodic characteristics of infant-directed speech as a function of maternal parity. In 9th International Conference on Speech Prosody 2018 (pp. 294–298).

[3] Beňuš, Š. (2014). Conversational entrainment in the use of discourse markers. In Recent Advances of Neural Network Models and Applications (pp. 345-352). Springer, Cham.

[4] Wang, T., Lee, Y. C., & Ma, Q. (2018). Within and Across-Language Comparison of Vocal Emotions in Mandarin and English. Applied Sciences, 8(12), 2629.

[5] Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. Cognition, 166, 84-93.

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

[7] Shue, Y.-L., P. Keating , C. Vicenik, K. Yu (2011) VoiceSauce: A program for voice analysis, Proceedings of the ICPhS XVII, 1846-1849.

[8] Garellek, M. (2019). The phonetics of voice. The Routledge handbook of phonetics, 75-106.

[9] Esposito, C., Sleeper, M., & Schäfer, K. (2019). Examining the relationship between vowel quality and voice quality. Journal of the International Phonetic Association, 1-32.

[10] Bartók, M. (2018). A gégeműködés variabilitása az érzelemkifejezés függvényében. Beszédkutatás, 26(1), 30-62.