

ChildASA dataset: Speech and Noise Material for Child-appropriate Paradigms on Auditory Selective Attention

Karin Loh  and Janina Fels 

Institute for Hearing Technology and Acoustics, RWTH Aachen University, Germany

Corresponding Author: janina.fels@akustik.rwth-aachen.de

Abstract: This dataset contains speech and noise material for a paradigm to investigate intentional switching of auditory selective attention in children. Two-syllable animal names were recorded from a male child and a female adult speaker. The spoken words were cut to the same length and adjusted to the same loudness to fulfill the requirements of a synchronous presentation. The noise material consists of a German four-talker-babble, recorded from children (all female), and a speech-shaped noise with the long-term spectrum of the children's four-talker-babble.

Keywords: Children, Stimuli Material, Speech, Noise, Auditory Selective Attention, German Animal Names

1 Introduction

To investigate intentional switching of auditory selective attention, Koch et al. [1] developed a paradigm including a categorization task in the presence of a target and a distractor speaker, which was extended to a binaural version by Oberem et al. [2]. The task hereby was to match the spoken stimulus from the target to the category bigger or lower than five. Thus, speech material comprising all digits from one to nine except zero and five were recorded. The respective datasets are downloadable in German via <https://doi.org/10.18154/RWTH-2020-02105> provided by Oberem and Fels [3] and in English via <https://doi.org/10.18154/RWTH-2020-08540> provided by Loh and Fels [4]).

To adapt this paradigm to children's knowledge starting from the age of three years old, the task was recreated with two-syllable animal names that can be categorized into the groups "flying" and "non-flying" by Loh et al. [5]. This paradigm further includes the background noise aspect besides the spatial sound environment allowing the investigation of noise effects on children's cognitive abilities. To represent children's daily environment more realistically contrary to normally in experiments used noises designed for adults, two noise types were designed: a multi-talker-babble (MTB) from four German girls comparable to the Swedish children multi-talker by von Lochow et al. [6] and a speech-shaped noise (SSN) with the long-term spectrum of the multi-talker-babble.

2 Speech material

The speech material consists of nine two-syllable animal names that were all phonetically dissimilar words: "Biene" (bee), "Ente" (duck), "Taube" (pigeon), "Eule" (owl) as flying animals versus "Katze" (cat), "Ratte" (rat), "Schlange" (snake), "Robbe" (seal) as non-flying animals in German (English translation in brackets). The dataset also includes a set of two-word combinations with the words "Kleine" (small) and "Große" (big) prior to the animal names. This two-word stimulus was designed to reach a longer stimulus length for future experiments with, e.g., room acoustics, where the acoustical effects might be longer than the stimulus itself. The spoken words were recorded by a male child and a female adult. All were native German speakers. The speakers were asked to use normal intonation and constant vocal effort as much as possible. Information about the speakers is listed in Table 1.

Table 1: Speaker's age, gender, and fundamental frequency

Speaker ID	Gender	Age (years)	Fundamental frequency (Hz)
mc1	male	5	209.6
wa1	female	26	240.2

3 Noise material

For the MTB, four German children (eight to nine years old), all female and native German speakers, were recorded as competing speakers. Each of the girls was



recorded individually while reading a short fairy tale from a children's book. The SSN is a filtered white noise with the long-term spectrum of the MTB.

4 Recording

Recordings were conducted using a diaphragm condenser microphone, Neumann TLM170 (cardioid directivity pattern at 24-bit resolution and 44.1 kHz), and a Zoom H6 hand-held recorder in the fully anechoic chamber of the Institute for Hearing Technology and Acoustic at RWTH Aachen University. Speakers were asked to position themselves approximately 1 m distance from the microphone.

5 Post-Processing

- All recordings were filtered using MATLAB¹ with a low-bandpass filter at 180 Hz.
- Recorded sequences for the speech material were cut to obtain the words. The desired maximum length for each word of the speech material was 800 ms. If the words were shorter than this, the necessary amount of silence was added in front and behind the word, situating the individual words in the middle of the sound file. Several recordings per word from every speaker were available for selection. The best version of the recorded stimuli was chosen based on subjective evaluations.
- Two-words combinations were generated by putting the words "Kleine" and "Große" before the animal name with 30 ms silence in between.
- Recorded sentences from the fairy tale were normalized for each girl and following put together on a four-track file to achieve a competing speaker pattern. All pauses longer than 300 ms and breath sounds were cut. The four-track file was constructed so that each girl's pauses almost always overlapped with other girls' speech to avoid unwanted pauses.
- All stimuli were equalized in loudness following the standard EBU-R 128 of the European Broadcasting Union using Adobe Audition².
- All files were exported to .wav with a 16-bit resolution.

6 Files

All files corresponding to this report can be downloaded from <https://doi.org/10.18154/RWTH-2023-00740>.

The dataset is sorted in the respective folders: Noises, Stimuli_oneWord, and Stimuli_twoWords. The stimuli folders further contain folders with the respective speaker id. Each file name is constructed following the pattern: "Speaker id"_"if existent: Große/Kleine"_"Animal name".wav. For example: *mc1_Biene.wav* or *wa1_Kleine_Katze.wav*.

Acknowledgments

The authors would like to thank all speakers for their participation in the recording. Special thanks to the student assistants Julian Burger, Julia Seitz, and Moritz Bender for assisting with recordings and post-processing. This work was funded by the HEAD-Genuit-Stiftung under project ID P-16/17-W and the European Union's Horizon 2020 research and innovation program under the grant agreement no. 874724 (Equal-Life). Furthermore, the project was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — project number: 368482240/GRK2416.

ORCID iD

Karin Loh  <https://orcid.org/0000-0002-5242-8784>

Janina Fels  <https://orcid.org/0000-0002-8694-7750>

References

- [1] I. Koch, V. Lawo, J. Fels, and M. Vorländer, "Switching in the cocktail party: Exploring intentional control of auditory selective attention," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 4, pp. 1140–1147, 2011. DOI: [10.1037/a0022189](https://doi.org/10.1037/a0022189).
- [2] J. Oberem, V. Lawo, I. Koch, and J. Fels, "Intentional Switching in Auditory Selective Attention : Exploring Different Binaural Reproduction Methods in an Anechoic Chamber," *Acta acustica united with Acustica*, vol. 100, no. 6, pp. 1139–1148, 2014. DOI: [10.3813/AAA.918793](https://doi.org/10.3813/AAA.918793).
- [3] J. Oberem and J. Fels, *Speech Material for a Paradigm on the Intentional Switching of Auditory Selective Attention*, 2020. DOI: [10.18154/RWTH-2020-02105](https://doi.org/10.18154/RWTH-2020-02105).
- [4] K. Loh and J. Fels, *English Speech Material for a Paradigm on Intentional Switching of Auditory Selective Attention*, 2020. DOI: [10.18154/RWTH-2020-08540](https://doi.org/10.18154/RWTH-2020-08540).
- [5] K. Loh, E. Fintor, S. Nolden, and J. Fels, "Children's intentional switching of auditory selective attention in spatial and noisy acoustic environments in comparison to adults," *Developmental psychology*, vol. 58, no. 1, pp. 69–82, 2022. DOI: [10.1037/dev0001239](https://doi.org/10.1037/dev0001239).
- [6] H. von Lochow, V. Lyberg-Åhlander, B. Sahlén, T. Kastberg, and K. J. Brännström, "The effect of voice quality and competing speakers in a passage comprehension task: Perceived effort in relation to cognitive functioning and performance in children with normal hearing," *Logopedics Phoniatrics Vocology*, vol. 43, no. 1, pp. 32–41, Jan. 2018. DOI: [10.1080/14015439.2017.1307446](https://doi.org/10.1080/14015439.2017.1307446).

¹ MATLAB is a software developed by MathWorks <https://de.mathworks.com/products/matlab.html>

² Adobe Audition is a digital audio workstation developed by Adobe Inc. <https://www.adobe.com/products/audition.html>