Design, structure, and preliminary analyses of a speech corpus of Infant Directed Speech (IDS) & Adult Directed Speech (ADS)

Katarzyna Klessa¹, Maciej Karpinski¹, & Agnieszka Czoska²
Adam Mickiewicz University in Poznan, Institute of Linguistics¹, Institute of Psychology²
(klessa, maciejk)@amu.edu.pl, agaczoska@gmail.com

Background

Infant-directed speech (IDS) is reported to differ significantly from adult-directed speech (ADS) in its acoustic, phonetic, and prosodic properties. In IDS, phonetic features of individual speech sounds tend to be intensified. For example, vowel hyperarticulation was documented for IDS in several languages. Consequently, the f0, formant frequency values (F1, F2) as well as timing and intensity patterns vary in the two speaking registers. Other modifications can be observed in pitch levels. Due to infants’ performance towards IDS, laboratory-elicited IDS recordings are often used as stimuli for the needs of infant speech perception studies, aiming e.g., at the investigation of the effects of short-term exposure to foreign language stimuli in early infancy and their role in the development of language learning skills. The latter often involves using speech samples from languages differing in phoneme inventories from the infant’s native language. In this paper, we present a corpus of IDS and ADS for the needs of phonetic training in infants. We also discuss the results of preliminary analyses by means of perception-based and phonetic-acoustic experiments.

Laboratory IDS & ADS

Languages & stimuli selection

- The corpus is expected to provide stimuli for training phonemic hearing in Polish infants, thus the choice of contrasts is primarily motivated by the presence / absence of specific phonemic contrasts in Polish.
- Illustrations of the IPA as a starting point: native speaker consultants.
- 5 languages, 2-5 speakers per language.

Recording scenarios

Each recording session consisted of two stages differing by stimuli elicitation manner:
- IDS mode (stage 1)
- ADS mode (stage 2)
In each mode speakers produced:
- isolated syllables,
- nonsense-words.

Stimuli presentation

Targets sounds for each contrast located in the same preceding / following contexts. In the nonsense-words, targets located in word-initial syllables but preferably not as word initial sounds.

Random order of presentation, 3 replications of each stimulus.

IDS mode

Speakers requested to speak in a manner as when addressing an infant.
Priming: a picture of a baby shown instantly before each stimulus.

Perception-based experiments

Question: same or different?

Test 1: French
153 pairs of syllables: vowel contrasts /e, æ/. Always preceded by /l/. Only one native speaker. 10 Polish native listeners advised to answer immediately after hearing the signal.

Test 2: Korean
153 pairs of syllables: consonant contrasts /p, b, pʰ/ (fortis unaspirated, lenis unaspirated, strongly aspirated). Always followed by /t/. 5 speakers. 20 Polish native listeners advised to answer immediately.

Test 3: French, Hungarian, Polish
56 syllables, 2 repetitions, random order. 6 speakers (2 per language). 16 Polish native listeners.
Continuous rating scale: Polish: foreign as shown in the Annotation Pro screenshot (right column of the poster). The middle of the scale = lack of any certainty. Listeners could re-play sound signals.

Question: native or foreign?

Phonetic-acoustic features

We compared pitch, F1, F2 values, and segmental duration in vowels produced by 5 female speakers of Polish, uttering pseudo-words in IDS and ADS. The results (Czoska et al., 2015) are as follows:
- systematic increase in segmental duration in IDS as compared to ADS,
- systematic increase in pitch in IDS,
- systematic increase in formant ratios in IDS.

Corpus applications

Apart from providing data for fundamental research, the present corpus currently serves as a source of stimuli in the electrophysiological (EEG) and eye-tracking-based studies of the development of the phonemic hearing and working memory in infants (see also the acknowledgements at the bottom of the poster).

Mother's speech

After the recordings at home the same mothers have been invited to participate in a recording session (first in ADS then in IDS) including:
- elitised speech: isolated & nonsense-words
- read speech: reading a passage from Turi’s Machines by St. Lem.

At the 48th Annual Meeting of the Societas Linguistica Europea, 2 – 5 September 2015, Leiden, the Netherlands.

Work supported by the project NeuroPaK: Development of the phonemic hearing and working memory in infants and children funded by National Science Center, Poland for years 2013-2016.

Selected references