Language Input to a Prelingual Infant

Joost van de Weijer

Max Planck Institute for Psycholinguistics
PO-Box 310
6500 AH Nijmegen
The Netherlands
vdweijer@mpi.nl

Abstract

Pitch, intonation, and speech rate were analyzed in a collection of everyday speech heard by one Dutch infant between the ages of six and nine months. Components of each of these variables were measured in the speech of three adult speakers (mother, father, baby-sitter) when they addressed the infant, and when they addressed another adult. The results are in line with previously reported findings which are usually based on laboratory or prearranged settings: infant-directed speech in a natural setting exhibits more pitch variation, a larger number of simple intonation contours, and slower speech rate than does adult-directed speech.

1. Introduction

The importance of language input cannot be doubted, since language development is clearly shaped by linguistic experience. This is very nicely shown in experimental work with infants that demonstrates that the infants lose the ability to discriminate non-native consonant contrasts in the second half of the first year of life (e.g., Werker & Polka, 1993; Werker, 1994). The influence of language input on language acquisition is also evident in experiments with prelingual infants between six and nine months (e.g., Jusczyk et al., 1993), which show that these infants exhibit a preference for the specific characteristics of their native language, such as the phonemic repertoire, prosodic regularities, and phonotactic constraints. In short: linguistic experience shapes speech perception. Children get attuned to the specific characteristics of their native language. This development takes place between the ages of six and nine months.

Soon after nine months children understand their first words (Benedict, 1979). In order to store words in the lexicon children have to segment words from the speech signal. However, the nature of spoken language poses a segmentation problem, because words do not usually occur in isolation but rather in context, and word boundaries are not reliably marked. This is true in adult-to-adult conversation, but also in adult-to-child conversation (Aslin, 1993). Thus children have to parse the input and decide which of the sound patterns are words, and which are not.

The fact that vocabulary building begins after the change in speech perception suggests that there is a link between this development and early word discovery, and that the period between six and nine months is crucial for early word segmentation.

Since this period is apparently very important for language acquisition it was decided to obtain a complete picture of the input that a child received between six and nine months. Therefore, almost all the input that an infant received was recorded. This was done in a Dutch family with two children (the infant and a sister two years older). Both parents had full time academic jobs, therefore the children were often taken care of by a baby-sitter or went to a day care centre. The language spoken was mainly Dutch. However, the mother, who was of German origin, often also spoke German to the children.

A DAT-recorder was carried around with the infant while she was awake, and care was taken that the cassettes were changed as necessary. The resulting collection of tape materials reflects about 90% of the time that the child was awake, a total of about 720 hours. Given the enormous amount of material, only speech from the first, middle, and last week of the period - 18 days in total - was included in the analysis. The language that was spoken during these days was extracted from the original recordings, stored on datatape, and analyzed with respect to a number of features.

In this paper I will focus on the suprasegmental structure of the language that the adults in this corpus addressed to the infant and will compare it with that of the language that the adults addressed to each other. The term *suprasegmental structure* refers to variations in pitch, timing, and amplitude of utterances. These variations can be determined by linguistic factors (such as: focal emphasis, syntactic structure, etc.) or by nonlinguistic factors (such as: different speech styles, different speakers, different emotions, etc.). In this paper, I focus on differences in suprasegmental structure that are the result of addressing either the infant or an adult.

The suprasegmental structure was analyzed because it has been previously shown that a distinct suprasegmental structure is one of the most salient characteristics of infant-directed speech (Fernald,

1989). Moreover, it has been shown that infants prefer to listen to speech in this speech style rather than to normal speech (Fernald & Kuhl, 1987). The variables that were measured are *pitch* (average pitch, pitch variation, pitch range), *intonation*, and *speech rate*.

2. Method

For this study utterances were selected from the three main adult speakers in the corpus: the mother, the father, and the baby-sitter. The choice of material was restricted because of the presence of frequent background noise, or overlap with other speakers. In general it was easier to select utterances to the infant than utterances to an adult. In the selection I aimed to choose utterances that were spread over the days, that had various lengths, and various conversational functions (e.g., questions, declaratives, discourse markers, etc.). An utterance was defined as a string that was a complete sentence, or a phrase that was delimited by a pause or an intonational bending indicating that this phrase was an utterance on its own.

All the material was digitized with a sample frequency of 16 kHz., and was processed with a speech editor. The beginning and end of each utterance was marked with the speech editor on the basis of visual inspection of the wave form in combination with auditory judgement. Pauses within utterances were also marked, and the duration of the pauses was subtracted from the total utterance duration. Furthermore, intonation contours were created for each utterance with the pitch extraction programme of the speech editor. Contours were excluded from the analysis when large parts were missing, or when incidental outliers fell outside the range of the rest of the contour, and could not be removed or corrected. Finally, the content of the utterances was transcribed orthographically.

For every utterance the following variables were determined:

— average pitch (Hz)
— standard deviation pitch (Hz)
— pitch range (Hz)
— pitch range (semitones)
— type of intonation contour
— utterance length (syllables)
— utterance duration (seconds)
— speech rate (syllables/sec.)

The intonation contours were classified according to their shape and their expansion of the frequency range (see: Fernald & Simon, 1984). The following types are distinguished: *bell-shaped*, *U-shaped*, *rising*, *falling*, *flat*, *complex* (multi-directional). These six contours have a simple shape or a wide frequency range (larger than six semitones per second) and are called *expanded* contours. If the contour did not meet these criteria, it was classified as *other*.

3. Results and Discussion

Table 1 presents the number of utterances that were selected from each speaker:

Table 1: Numbers of selected utterances.

	to infant	to adult
mother	734	874
father	574	1113
baby-sitter	575	506
total	1883	2493

The numbers vary because the amount of material available from each speaker varied, and because not all the material from each speaker could be used equally well (e.g., a relatively large number of utterances produced by the father could not be used for the pitch analysis). For the measurement of pitch and intonation 1304 utterances (29.8%) were not used because the intonation contours were not reliable. The reason why a relatively large number of utterances had to be discarded was that these recordings were not made under ideal conditions. The microphone was always close to the infant but the speakers moved freely through the room. This often resulted in a low signal-to-noise ratio, especially in the adult-directed speech. The infant-directed speech which was usually produced close to the microphone was less affected. The results of the pitch measurements are presented in table 2.

Table 2: Pitch measurements for infant-directed speech (first values) and adult-directed speech.

	mother	father	baby-sitter
average (Hz)	279 — 233	157 — 130	303 — 225
minimum (Hz)	228 — 196	126 — 111	234 — 180
maximum (Hz)	349 — 291	190 — 157	386 — 276
range (semitones)	7.22 — 6.64	7.14 — 5.80	8.89 — 7.27
stand.dev. (Hz)	36 — 25	19 — 13	47 — 28

Table 3: Percentages of occurrence of contour types in infant-directed speech (first values) and adult-directed speech.

	mother	father	baby-sitter
rising	19.0 — 10.0	24.5 — 20.9	24.3 — 11.1
falling	10.6 — 18.9	10.3 — 20.7	13.5 — 12.5
bell-shaped	11.6 — 3.1	13.1 — 0.2	8.7 — 5.8
U-shaped	0.8 — 2.2	0.4 — 0.7	3.5 — 1.5
level	3.2 — 0.0	3.5 — 0.0	10.4 — 0.0
complex	29.4 — 22.4	27.7 — 23.1	12.2 — 19.8
other	25.4 — 43.4	20.6 — 34.3	27.4 — 49.3

It can be seen from the table that across the speakers, the infant-directed speech has higher average pitch, higher minimum and maximum pitch values, wider range and higher standard deviation. These findings are in line with previously reported findings in other languages, such as English, Japanese, German, French, Chinese (e.g., Fernald et al., 1989; Grieser & Kuhl, 1988).

Table 3 shows the distribution of the contour types in the utterances that were also used for the pitch measurements. The percentages of expanded contours (the sum of the first six types) are higher in infant-directed speech than in adult-directed speech. The differences are not as clear as in Fernald & Simon's (1984) study. They found a very low percentage of expanded contours in adult-directed speech and therefore concluded that these expanded contours are typical of infant-directed speech. The reason why in this study I found a relatively large number of expanded contours in adult-directed speech is that the choice of material was restricted to relatively short utterances. For that reason, many of the adult-directed utterances did meet the six-semitones-per-second criterion, and could not be classified as *other*.

Most of the selected utterances could be used for the temporal measurements. A total of 204 utterances (4.7%) were discarded because the exact number of syllables could not be determined. The average duration (total duration minus pause duration) of the infant-directed utterances was 1.01 sec., and the average duration of the adult-directed utterances was 0.89 sec. The average length of the infant-directed utterances was 3.32 syllables, and the average length of the adult-directed utterances was 5.64 syllables. Previous studies have also reported relatively short utterances in infant-directed speech (e.g., Phillips, 1973).

However, the difference in duration seems to be relatively small compared to the difference in length, which suggests that the adult-directed speech was produced with a higher speech rate. The measurements of speech rate are shown in table 4. The values are broken down by utterance length since shorter utterances tend to have a slower speech rate. As can be seen from the table the utterances to the infant are produced at a slower rate than the utterances to the adults. The values are similar to those reported by Fernald & Simon (1984), who found an average speech rate of 4.2 syllables per second in infant-directed speech, and 5.8 syllables per second in adult-directed speech.

Table 4: Speech rate broken down by utterance length (syllables) of infant-directed speech (first values) and adult-directed speech.

	mother	father	baby-sitter
1	2.44 — 3.33	2.88 — 4.25	2.07 — 3.00
2	3.40 — 4.58	3.83 — 5.83	2.66 — 3.95
3	3.88 — 5.47	4.19 — 5.46	2.88 — 3.89
4	4.45 — 5.66	5.00 — 5.82	3.56 — 4.22
5	4.84 — 6.15	5.44 — 6.30	3.78 — 4.76
6	5.06 — 6.56	5.07 — 6.69	4.33 — 5.15
7 or more	5.76 — 6.48	6.56 — 6.98	4.71 — 6.02
average	3.81 — 5.27	4.10 — 5.72	3.16 — 4.64

4. Conclusions

The results of this study show that the suprasegmental structure of the infant-directed speech is quite different from that of the adult-directed speech. This result has been reported before, but was always based on data obtained in prearranged settings, often in laboratories. The material for this study was part of a larger collection of every day speech in the child's own situation. We measured pitch, intonation, and speech rate in the speech of the speakers from which this infant received most of her input. The differences in suprasegmental structure may facilitate the discovery of the structure of the language in the initial stages of language acquisition: The high pitch and the large pitch variations serve to attract and maintain the infant's attention; the simple intonation contours make utterance boundaries more clear; the slow speech rate gives the child more time to process the speech signal.

This study was done as a first step in finding out more about the relation between input and acquisition. Further analyses will follow, focusing more directly on word discovery.

References

Aslin, R. (1993). Segmentation of fluent speech into words: learning models and the role of maternal input. In: B. Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage & J. Morton (Eds.), *Developmental Neuro-cognition: Speech and Face Processing in the First Year of Life*. Kluwer Academic Publishers, Dordrecht, Boston, London.

Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language*, **6**, 183-200.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, **92**, 81-104.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Development*, **60**, 1497-1510.

Fernald, A. & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, **10**, 279-293.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B. & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, **16**, 477-501.

Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, **20**, 104-113.

Grieser, D. & Kuhl, P. (1988). Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Developmental Psychology*, **24**, 14-20.

Jusczyk, P., Cutler, A. & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, **64**, 675-687.

Phillips, J. (1973). Syntax and vocabulary of mothers' speech to young children: age and sex comparisons. *Child Development*, **44**, 182-185.

Werker, J. (1994). Cross-language speech perception: developmental change does not involve loss. In: J. Goodman & H. Nusbaum (eds.). *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. The MIT Press, Cambridge, Massachusetts, London, England.

Werker, J. & Polka, L. (1993). Developmental changes in speech perception: new challenges and new directions. *Journal of Phonetics*, **21**, 83-101.