

VOICE ONSET TIME IN SPEECH DIRECTED TO INFANTS AND ADULTS

Ulla Sundberg

Department of Linguistics, Stockholm University

ABSTRACT

Voice Onset Time (VOT) was measured in three Swedish mothers' infant-directed speech (IDS) and compared to VOT in their adult-directed speech (ADS). In this preliminary study VOT was significantly shorter in IDS than in ADS. The impact of stress was very clear in both IDS and ADS showing significantly longer VOT in stressed positions as compared to the unstressed. The shorter VOT in IDS and the finding that the absolute range of VOT values was smaller in IDS than in ADS, may suggest vocal accommodation [1] from the mothers' part.

INTRODUCTION

Infant-directed speech (IDS) differs in several aspects from adult-directed speech (ADS). The most striking differences are the well documented prosodic modifications with exaggerated intonation contours, shorter utterances and longer pauses, for example [1]. Acoustic investigations of IDS and ADS at the segmental level are rather scarce although some studies have been performed regarding vowel formant structure [2] and vowel modification [3] and pre-boundary vowel lengthening [4] e.g. IDS is often referred to as being more clear and informative than ADS [3, 5, 6]. One factor that might influence clearness in speech is Voice Onset Time (VOT), i.e. the duration of the gap between the burst marking the release of the articulatory closure and the onset of voicing in stops. In most languages

VOT may serve as one factor separating voiced from voiceless segments [7, 8] into the two phonemic categories. In speech directed to children during their early language acquisition period, it was found that, in contrast to ADS where VOT varied drastically in duration, the mothers prolonged VOT in voiceless stops, thus producing extreme ones, and in this way efficiently separated voiceless stops from their voiced cognates, (Bernstein-Ratner, 1984, [4] referring to an unpublished investigation by Moslin, 1979.)

One other aspect of IDS that could be illustrated by analysis of VOT is phonetic accommodation, i.e. the interactors' mutual affective engagement resulting in accommodation of speech style to one another [1]. One expression of intimacy in the interaction between the mother and her baby might be reflected in the mother's way of mimicking young children's early speech with e.g. unaspirated stops [9]¹.

The objective of the present study is to make a preliminary assessment of how Swedish mothers' use VOT in IDS and ADS.

PROCEDURE

Three Swedish mothers interacted with their three-month old infants, two girls and a boy, in an isolated booth for 10 to 15 minutes. The mothers were instructed to play with some toys in a way they would normally do at home. The IDS sample was collected from this

¹ Older children at the age of 3-4 years may produce longer VOT than adults [10].

session. The ADS sample was collected immediately after this session. The investigator entered the room and talked informally with the mother about different topics concerning the infant. The sessions lasted 20-30 min. The first 4-5 minutes of the IDS and ADS samples were selected for analysis. Before analysing the data the investigator made an auditory judgement of the dialogues marking the words that were perceived as being the most prominent one(s) in each utterance. VOT in the syllables carrying lexical stress in these prominent words are in the following called stressed. The VOT was defined as the time gap between the onset of the burst and the onset of voicing, both marked on the speech wave signal.

RESULTS

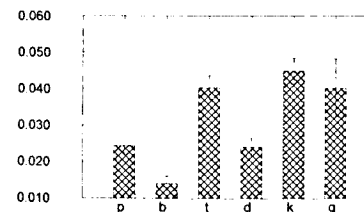


Figure 1a. Mean VOT in msec, in IDS. The bars indicate the standard error.

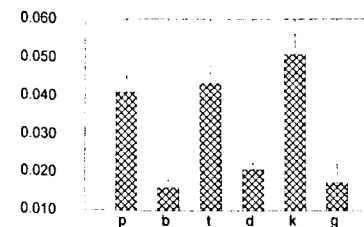


Figure 1b. Mean VOT in msec, in ADS.

VOT values in stops in word initial and medial positions were subjected to a

3-way analysis of variance. Addressee, consonant type and stress were the three factors. All three mothers showed great overlap of VOT within the categories voiced and voiceless in both speech directions. VOT was significantly longer in ADS as compared to IDS, ($F(1,222)=3.955$, $p < 0.05$). A very significant difference was found regarding consonant type and VOT duration, ($F(2,222)=6.812$, $p < 0.001$), showing that the labials had shorter VOT than the dentals, and the dentals had shorter than the velars, see figure 1a and b. The voiced stops were excluded from further analysis due to uneven numbers or missing values in some cells, which in turn reflects the finding made in earlier investigations e.g. [7] that Swedish voiced stops lack aspiration phase.

The results revealed significantly longer VOT in words with stress than in words without stress ($F(1,222)=9.656$, $p < 0.002$). No interaction was found between the two factors of stress and consonant type ($F(2,222)=0.326$, $p > 0.7$). A strong interaction, on the other hand, was found between stress and addressee ($F(1,222)=3.128$, $p < 0.08$).

DISCUSSION

Adults use a speech style qualitatively different in their interaction with infants as the style that is used amongst adults. The mothers in this investigation produce shorter VOT in IDS, regardless of stress, than in their ADS. These results may indicate that mothers adjust their speech to a style closer to children's own early speech. (Phonological reductions in IDS may serve a similar purpose as suggested by Shockey & Bond (1980), [11], by "setting a tone of intimacy in a dialogue"). Vocal accommodation have mainly been suggested in the context of children's vocal behaviour [1]. There

are no obvious reasons though, to believe that adults wouldn't adjust their speech to closer conformity with children's speech when interacting with infants.

The influence of stress on VOT was shown very clearly in this study by the significantly longer values in stressed positions as compared to unstressed positions. These results are in accordance with earlier investigations of both so called lab speech [7, 12] citation form, [13, 14] and spontaneous speech [14]. Krull (1991) who made the first VOT measurements of VOT in spontaneous Swedish found 30-100% longer VOT in stressed CV-sequences than in corresponding unstressed syllables. A noteworthy aspect in the present study is that the range of the VOT values in IDS was smaller than in ADS, suggesting a less pronounced separation of voiced and voiceless stops in IDS by means of VOT. This is contrary to the expectations, but it must be kept in mind though that VOT is only one out of several acoustic features affecting the voiced/voiceless distinction. Löfqvist (1976), [12] e.g. suggests a complex differentiation of voiced and unvoiced segments in terms of closure duration, which is inversely related to VOT, comprising acoustical duration of the consonant, acoustical duration of the vowel, duration and spectral extensiveness of the vowel formant transitions, to mention some.

The narrower range of VOT values in absolute terms in IDS as compared to ADS lends support to the suggestion mentioned earlier that VOT in IDS may be one phonetic feature signalling the mothers' vocal accommodation [1]. At the age of 11 months children's production of stop-vowel syllables is often characterised by simultaneous, or almost simultaneous, release of the stop and onset of voicing [9]. By using a

speech style mimicking children's way of speaking, virtual 'Baby Talk', the mother adjusts her speech into closer conformity with her interlocutor, a phenomenon that according to Giles (1984) may unconsciously be increasing the perceived attractiveness to the listener and the level of involvement.

The discrepancy between the results from the present investigation showing shorter VOT in IDS than in ADS and those from Moslin (1979), [4], who found the inverse relationship are striking. A possible explanation may be a developmental change of the characteristics in IDS related to the age and language development of the addressee [15]. In Moslin's investigation the children were "in the first stages of language learning", i.e. in the age range 12 to 20 months. The function of IDS to children of these ages are often related to object play and language acquisition [15] that may affect a phonetic feature such as VOT to have a more definitive clarification function. The age of the addressees in the present study is only three months, and the social context and the kind of interaction is here very different from that mentioned in the previous study. In the interaction with young infants the goal is to engage them in a face-to-face play, the tone is intimate and in the present study the mothers often encouraged the infants to vocalise. The differences in context and in goal of interaction in the two studies may thus account for the discrepancy of VOT in the observed IDS and ADS patterns.

ACKNOWLEDGEMENT

I would like to express my gratitude to Francisco Lacerda for assistance with the statistical analysis and helpful discussions.

REFERENCES

- [1] Locke, J.L. (1993), *The child's path to spoken language*, Harvard University Press, Cambridge, Mass.
- [2] Davis, B. & Lindblom, B. (1992), "Prototypical vowel information in Baby Talk", *PERILUS*, No XV, pp 119-124.
- [3] Bernstein Ratner, N. (1984), "Patterns of vowel modification in mother-child speech", *Journal of Child Language*, 11, pp 557-578.
- Moslin, B. (1979), "The role of phonetic input in the child's acquisition of the voiced-voiceless contrasts in English stops: A VOT analysis", Unpublished doctoral dissertation, Brown University.
- [4] Bernstein Ratner, N. (1986), "Durational cues which mark clause boundaries in mother-child speech", *Journal of Phonetics*, 14, pp 303-309.
- [5] Fernald, A. (1983), "The perceptual and affective salience of mothers' speech to infants", in *The Origins and Growth of Communication*, (Eds. L. Feagans, C. Garvey, R. Golinkoff) Ablex Publishing Corporation, Norwood, New Jersey, pp 5-29.
- [6] Fernald, A. & Mazzie, C. (1991), "Prosody and focus in speech to infants and adults", *Developmental Psychology*, Vol. 27, No. 2, pp 209-221.
- [7] Fant, G. (1973), *Speech sounds and features*, MIT Press, Cambridge, Mass. p 110 ff.
- [8] Lisker, L. & Abramson, A., (1964), "A cross-language study of voicing in initial stops: Acoustic measurements", *Word*, 20, pp 384-422.
- [9] Fletcher, P. & Garman, M. (1986), *Language Acquisition, Studies in first language acquisition*. Cambridge University Press, p. 161.
- [10] Menyuk, P. & Klatt, M. (1974), "Voice onset time in consonant cluster production by children and adults", *Journal of Child Language*, 2, pp 223-231.
- [11] Shockey, L. & Bond, Z.S. (1980), Phonological processes in speech addressed to children. *Phonetica* 37: 267-274.
- [12] Löfqvist, A. (1976), "Closure duration and aspiration for Swedish stops", *Working Papers 13, Phonetics Laboratory, Dept of General Linguistics*, Lund University, pp 1-39.
- [13] Engstrand, O. (1983), Articulatory co-ordination in selected VCV utterances: A means - end view. (Ph.D. diss.) *Reports from Uppsala University Department of Linguistics* (RUUL) 10.
- [14] Krull, D. (1991), "VOT in spontaneous speech and in citation form words", *PERILUS* XII, Stockholm University, pp 101-107.
- [15] Stern, D.N., Spieker, S, Barnett, R.K., & MacKain, K. (1983), "The prosody of maternal speech: Infant age and context related changes", *Journal of Child Language*, 10, pp 1-15.