

A Study of Voicing in the Stops Found in the Pre-Linguistic
Babblings of Infants

Malcolm S. Preston[†]

Haskins Laboratories, New York City

Starting at about six months of age, the frequency and diversity of speech-like sounds in the vocalizations of infants with normal hearing show a rapid increase(1). By the age of twelve months, a child may have several different sounds which appear periodically in his babblings. Quite often, a mother will report that her child said "da da" at ten or twelve months although it is doubtful that many children could use such an expression in a meaningful communicative sense at that age.

Much of the work on the acquisition of the phonological system has been done by researchers who relied chiefly on traditional methods of linguistic analysis. Recent work at Haskins Laboratories employing both spectrographic analysis and synthetic speech has made possible a new approach to this problem. A study of the acoustic correlates of stops in initial position produced by speakers of eleven different languages(2) suggests that the time interval between the release of air pressure and the onset of voicing is an effective measure for sorting the stop consonants into phonemic categories. Although boundaries between contrasting categories differ from language to language, it appears that three general positions along the continuum of voice

onset time may be utilized: (i) 70-100 msec. voicing lead, (ii) 0-30 msec. voicing lag, and (iii) 50-110 msec. voicing lag. Voicing lead in the 0-40 msec. range was rarely noted in any of the eleven languages.

While some languages like Thai make use of all three positions, others use only selected portions along the continuum. For instance, English and Cantonese Chinese make use of the two voicing lag positions while Spanish and Russian make use of the 70-110 msec. lead and 0-30 msec. lag positions. Thus for apical stops, the Spanish /t/ and the English /d/ both fall in the 0-30 msec. lag position while the English /t/ falls in the 50-110 msec. lag position and the Spanish /d/ falls in the 70-110 msec. lead position.

Since not all languages employ the same positions along the voice onset time (VOT) continuum, a cross language study of voicing in stops found in the pre-linguistic babblings of infants would cast some light on the development of this important distinctive feature. For instance, a longitudinal study comparing the pre-linguistic stops of English with Spanish infants would determine at what point a child begins to utilize voicing as a means of distinguishing between two classes of phonemes (/b,d,g/ vs /p,t,k/). One possibility is that all children, irrespective of their linguistic environment, will exhibit the same range of VOT values when stops in initial position are first produced. Another possibility is that each child will exhibit, from the start, VOT values peculiar to the language of his parents or caretaker. The present study is a first step towards examining these and other possibilities. This research is being carried on in collaboration with the Neurocommunications Laboratory, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine in Baltimore, Md.

Method

Some of our preliminary investigations with infants have shown that not long after nine or ten months of age, normal hearing children will begin to produce in their babblings recurrent vocalizations with stop-like qualities. Prior to this age, that the neural structures and other physiological factors underlying the production of these sounds are still in a very primitive stage of development. It seemed reasonable, then, to start collecting vocalization data from infants at ten months of age, and to continue collection until at least 24 months.

Some longitudinal vocalization data were already available from a study initiated by the late Ruth Weir at Stanford University. These recordings were made two or three times a month from about the age of six months. The parents of these children were provided with a tape recorder and instructed to turn the machine on when the child was vocalizing. There were some English, Russian and Chinese infants in the sample. Unfortunately, the recording conditions were very poor, making most of these data unanalyzable on the spectrograph. Some of the data from one of the Chinese infants were usable and are reported in this paper.

Arrangements were made to collect data from six children of the medical staff at Johns Hopkins Hospital. Recordings were made on a weekly or biweekly basis, whichever was most convenient for the mother. When recordings started, four children were in their tenth month and two were in their twelfth month. There were three males and three females. All children came from American English speaking families and had at least one older sibling.

Mother and child were brought to the laboratory and placed in a small double-walled IAC booth furnished with several small chairs and a few toys. A Bruel and Kjaer type 4131 one-inch condenser microphone was suspended three feet from the floor in the center of the room. The microphone was connected to an Ampex model AG 350 tape recorder. The mother was instructed to inter-

act verbally with the child in her usual manner. She was cautioned not to talk while the child was vocalizing. One or two 25-minute recording sessions were made for each visit to the laboratory.

The mothers were told that the research was being carried on to examine the sorts of sounds very young children make as they become older. In addition to the six children described above, other children in the 10 to 16 month range have been brought to the laboratory for recording sessions when the opportunity arose. Data collected from a child being raised in a bilingual English-Hungarian family are reported in this paper. Since Hungarian (like Spanish) uses the position of voicing lead as well as the 0-30 msec. lag position on the VOT continuum, and English uses the 50-110 msec. lag position, it is reasonable to suppose that this child is being exposed to stops from all three categories.

Results

Tape recordings of the regular series of Hopkins infants are still being collected. Results for one recording session are presented below for four infants including two from the regular series. Two of these children (E1 and E2) have American English speaking parents, one child (C) has Cantonese speaking parents and the fourth child (H) comes from a bilingual English-Hungarian family. The ages at the time of recording were as follows:

E1: 12 months, 1 week

E2: 13 months, 1 week

C: 14 months, 1 week

H: 12 months, 3 weeks

To begin the analysis, a judge (MSP) listened over Sharpe headphones to each recording session and re-recorded any sounds which appeared to have a stop-like quality. The edited tapes were then presented over headphones to two other judges who were asked (1) to eliminate any sounds they felt were not stops

and (2) to identify the place of closure (labial, apical or velar) for sounds judged to be stops. Wide-band spectrograms (600 cps band-width) were made using a Voice Print, Inc. spectrograph. Only sounds judged to be stops by all three judges and on which there was 100% agreement on place of closure were analyzed.

To determine VOT, measurements were taken directly from the spectrograms. Measurements were not taken if the onset of voicing or the release of air pressure were not both clearly evident, or if voicing was continuous from a preceding vowel sound. The latter case was more likely to occur if the time interval between "syllables" was short (less than 100 msec.). From a total of 374 items for all four subjects, 91 (24%) were rejected because of unmeasurable spectrograms and 58 (15%) because of disagreement among the judges. The figure presents frequency distributions for VOT for labial, apical and velar stops respectively. The data have been grouped in 5 msec. intervals. No child produced a sufficient number of stops to warrant making frequency distributions for all three positions of closure. Subject E1 produced no labials while Subject H produced only two labials. Subject E2 produced no apicals. Subject H produced 7 velars with VOT ranging between 210 msec. voicing lead and 70 msec. voicing lag while Subject E1 produced 8 velars ranging between 80 msec. lead and 75 msec. lag. Subject C produced no velars.

Discussion

In contrast to the bi-modal patterns usually found in the frequency distributions of VOT in adult speakers of English, Cantonese and Hungarian, the patterns shown in the figure are clearly uni-modal. In general, the majority of stops fell between 25 msec. lead and 25 msec. lag, although for Subject H, a considerable number of stops show long voicing lead. For all

subjects, there were relatively few stops with voicing lag greater than 30 msec. Thus, although more data are necessary, something approximating the middle category on the VOT continuum is the first to emerge in infant vocalizations. In addition, it appears that a child may learn to produce stops with long voicing lead before he learns to produce stops with long voicing lag. One other point worth noting in these preliminary data is a tendency for velar stops to have the greatest VOT values followed by apical stops and labial stops. This is clearly evident in comparing labial with velar stops for Subject E2, and labial with apical stops for Subject C. A similar finding has been noted for adult speakers(2).

To supplement the analysis of infant vocalizations, an analysis is underway to determine VOT in the stops produced by the mother while interacting with her child. In addition, an opportunity has arisen to collect data from several 12 to 14 month old infants being raised in a Lebanese Arabic speaking environment. These data will be available shortly and will provide an interesting comparison with the data in this report since Arabic employs the long voicing lead and short voicing lag positions. Finally, a study of the vocalizations of severely deaf children before and after obtaining a hearing aid is being carried out currently at Johns Hopkins University Medical School and will provide additional data for comparison.

Footnotes

† Also, Neurocommunications Laboratory, the Johns Hopkins University School of Medicine.

1. Irwin, O.C. and Chen, H.P. Infant speech: vowel and consonant frequency. *J. Speech Disorders* 11, 123-125 (1946).
2. Lisker, L. and Abramson, A.S. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20, 384-422 (1964).