A Study of Voicing in Initial Stops Found in the Pre-Linguistic Vocalizations of Infants from Different Language Environments*.

Malcolm S. Preston

Haskins Laboratories, New York City
Grace Yeni-Komshian and Rachel E. Stark
Neurocommunications Laboratory, The Johns Hopkins University
School of Medicine

Introduction

Our research has been concerned with exploring the effects of the language environment on the development of speech motor capabilities. We have been particularly interested in how a child gains control over his vocal apparatus such that his vocalizations contain acoustic features which are critical for speech communication. In pursuing our studies, we have made comparisons between the vocalizations of children raised in two different language communities (English and Lebanese Arabic), paying close attention to acoustic features known to differ in the two languages.

In a study here at Haskins Laboratories (1), spectrographic analysis of stop-consonants produced by speakers from 11 language communities suggested that the time interval between the release

* Given at the 73rd Meeting of the Acoustical Society of America, New York City, 19-22 April 1967.
† Also, Neurocommunications Laboratory, The Johns Hopkins University School of Medicine.
of air pressure and the onset of voicing was a useful measure for sorting word-initial stop-consonants into phonemic categories. Voice onset time (VOT) was obtained directly from wide-band spectrograms by measuring the distance between the first vertical striation representing glottal pulsation and the onset of energy in the formant frequency range representing the release of air pressure. In some cases, voicing actually commenced during the period of occlusion preceding the release of air pressure. When the distributions for all eleven languages were combined, three general ranges along the VOT continuum were noted: (i) 50-150 msec voicing preceding release of air pressure (lead), (ii) 0-30 msec voicing following release (short lag), and (iii) 50-110 msec voicing following release (long lag).

Figure 1 shows spectrograms of apical stops produced by adult speakers of English and Arabic, and illustrates typical VOT values for the two languages. The VOT values for the two English words at the top are 50 msec voicing lag for "tapper" and 0 msec for "dapper." The two Arabic words at the bottom, "taraka" and "darasa," have VOT values for 15 msec voicing lag and 65 msec voicing lead respectively.

Figure 2 shows VOT distributions for speakers of English, from the work of Lisker and Abramson (1). This figure shows that /b, d, g/ fall mainly between 0-30 msec voicing lag while /p, t, k/ fall between 40 and 100 msec voicing lag. Figure 3 shows distributions of data we have recently collected from adult speakers of Arabic (2). This figure clearly indicates the incidence of voicing lead (30-150 msec) for /b, d/ and emphatic /d/. The other stops /t/, emphatic /t/, /k/ and /q/ fall mainly between 0-50 msec voicing lag.
Words in isolation for adult speakers of English


SLIDE 2
WORDS IN ISOLATION FOR ADULT SPEAKERS OF LEBANESE ARABIC
SLIDE 3
In summary, in American English measurement of VOT can be used to distinguish /b/ from /p/, /d/ from /t/, and /g/ from /k/. The phonemes /b, d, g/ utilize the short lag position and to a lesser extent the lead position, while /p, t, k/ utilize the long lag position. On the other hand, the Lebanese data demonstrate that Arabic makes use of the voicing lead and short voicing lag positions for distinguishing /d/ from /t/, /d/ utilizing voicing lead and /t/ short voicing lag. The emphatic apicals, /d/ and /t/, can be distinguished in the same fashion. In addition, Arabic /b/ is generally produced with voicing lead while the two velar stops, /k/ and /q/, utilize short voicing lag. Lebanese Arabic does not have /p/ or /g/.

The children of both American English and Lebanese Arabic speaking parents are exposed to stops with VOT values falling in the short voicing lag position. However, American Children are exposed to stops with long voicing lag while Lebanese children are exposed to stops with voicing lead. Since the linguistic-acoustic environment with respect to VOT differs for children from these two language communities, a cross-language study of voicing in stops occurring in the pre-linguistic vocalizations of these children may cast some light on the development of this important distinctive feature. One possibility is that children from both language communities will exhibit the same range of VOT values before communicative language skills are acquired. Another possibility is that, at this early stage of speech development, each child will exhibit VOT values peculiar to the language of his parents.

We will report on the vocalizations of four American and three Lebanese children ranging in age from 50 to 57 weeks. In
addition, we will present data from the vocalizations of two hearing impaired American children (104 and 106 weeks) who were examined between 26 and 29 weeks following the acquisition of a hearing aid. Interest in these last two cases centered on determining whether the deaf children who had been deprived of normal hearing up to the age of 77 or 78 weeks would exhibit VOT values similar to the younger normal hearing children from the same language community. The American children are part of a larger longitudinal study.

Method

The vocalizations of the children were tape recorded over a period which lasted between 30 and 60 minutes. The mother and sometimes an experimenter were present. The child was encouraged to vocalize as much as possible.

In the analysis, a judge listened over headphones to each recording session and noted all sounds which appeared to have a stop-like quality occurring in initial position in a syllable. 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 point of articulation (labial, apical or velar) for sounds judged to be stops. Only sounds classified as stops by all judges and on which there was complete agreement on point of articulation were analyzed spectrographically. To determine VOT, measurements (to the nearest 5 msec) were taken directly from wide band spectrograms.

Frequency distributions for each child were constructed for each point of articulation if there were at least ten observations.
The results to be presented now for the normal hearing children are based on one recording session for each child at about one year of age.

Results and Discussion

Figure 4 shows one labial, two apical and two velar distributions for the normal hearing American children. In contrast to the bimodal distributions seen for the adults, these distributions are essentially uni-modal. The distribution for labial stops shows a greater incidence of voicing lead than that noted for adults. However, the distributions for the apical and velar stops closely match the adult models for /d/ and /g/.

Figure 5 shows one labial, three apical and one velar VOT distribution for the Lebanese children. Although these children are exposed to different VOT models, their results are very similar to the American children. The distribution for labial stops shows voicing lead, while the distributions for apical and velar stops show voicing lag mainly between 0 and 30 msec.

Figure 6 shows two labial, one apical and one velar VOT distribution for two deaf children about 2 years of age who had been wearing hearing aids for about 26 weeks. These children did not produce any stops before acquiring a hearing aid. Again, these distributions are not too unlike those for the younger normal hearing children. They are essentially uni-modal. The labial stops show voicing lead although their modes are at 0 msec. The apical and velar stops fall mainly between 0 and 30 msec voicing lag.

The following points summarize the results. First, the
SLIDE 5

LEBANESE CHILDREN: NORMAL HEARING

LABIAL
A6  N=21
57 WEEKS

APICAL
A3  N=23
50 WEEKS

APICAL
A6  N=38
57 WEEKS

APICAL
A8  N=12
50 WEEKS

VELAR
A8  N=36
50 WEEKS

LEAD -200 -150 -100 -50 0 50 100
VOICE ONSET TIME IN msec
LAG
distributions for normal hearing American and Lebanese children are quite similar. With respect to apical stops for which the most data is available, these children show uni-modal distributions falling mainly in the 0 to 30 msec voicing lag range. This occurs in spite of the fact that the Lebanese children had been exposed to apical stops with voicing lead while the American children had been exposed to apical stops with voicing lag between 40 and 100 msec or greater.

Second, the older hearing-loss children showed VOT distributions similar to the younger normal children. This suggests, perhaps, that deaf children with hearing aids may acquire the ability to produce the two classes of stops /b, d, g/ and /p, t, k/ in the same manner as the normal child. Our longitudinal data will help to clarify this possibility.

References
