CLUSTERS AS SINGLE UNDERLYING CONSONANTS: EVIDENCE FROM CHILDREN'S PRODUCTION*

PAULA MENYUK

Some historical data (Gamkrelidze 1966) indicates that some clusters are equivalent to a single consonant at some point in time and then later are realized as two separate sounds. It has also been proposed (Hoffman 1967) that initial clusters are best described as a single element in the underlying phonological representation. Later rules decompose the cluster into segments with each segment containing its complete set of features. The underlying representation of a speech sound and of the clusters in which it participates would be different, as indicated in Figure 1.

	k	k_r	k_l	k_w
Obstruent	+	+	+	+
Consonantal	+	+	+	+
Rounded	–	+	–	+
Coronal	–	+	+	–

Fig. 1. Underlying representation of singleton consonant and possible clusters.

How initial clusters are stored in the human perceiver's lexicon is not specified by the above descriptions. The way in which the child goes about acquiring initial clusters, either segment by segment or as a single consonant with special features, can provide some evidence of the way in which clusters are perceived and stored in the lexicon before the complete set of rules for cluster generation is acquired and used. This information may have some bearing on the question of what the basic perceptual unit is. If clusters are perceived as some distinctive features of the first consonant, etc., then one might hypothesize that the basic unit is the distinctive feature or small subsets of features. If clusters are perceived as consonant plus consonant, then perhaps the basic unit is the speech sound segment or phonemic categorization. Finally, if clusters are perceived as a single underlying consonant, then one might hypothesize that the basic perceptual unit is the syllable, since in this last instance the perceiver maintains the consonant plus vowel structure of the syllable.

* Read by A.W.F. Huggins.

The utterances of the child (previously recorded by Bullowa *et al.* 1964) were sampled at six month periods from 18 months to age 30 months and again at 34 months. Utterances which contained initial clusters were identified and re-recorded for analysis. These recordings were phonetically transcribed by four experienced listeners, spectrograms were made of each utterance and tracings of formant transitions were made from these spectrograms. In addition, spectrograms were made of utterances that contained the same singleton initial consonant and vowel as those found in some cluster utterances (for example, Brian and bike, blow and boat). Only a small set of these minimal contrast pairs could be found in the language sample. A computer-controlled filter-bank spectral analyzer was used to supplement the spectrographic analysis of these pairs. The techniques used are reported in Menuyk and Klatt (1968).

It was found that consonant + /l/, /w/ and /r/ were initially produced as a singleton consonant but that this singleton consonant of a cluster was different than that found in similar utterances with no initial cluster. This difference was primarily one of duration. The transitions to the peak of the vowel were longer in all instances of the cluster than in the singleton consonants. This aspect of gradual release was later realized homorganically in some instances so that 'truck' was produced as /chuck/ and 'driveway' was produced as /jiveway/. Clusters with /s/ + consonant were initially produced as ø + consonant in the case of stops and nasals (/guy/ for 'sky' and /neager/ for 'sneaker') but as s + ø in the case of /r/ and /l/ (/sim/ for 'swim' and /seep/ for 'sleep'). Kornfeld (1971) has found that even though the /s/ is omitted in /s/ + stop clusters, the consonant is marked in some way for stridency and gradual release. Thus, not only are difference observed between singleton consonants and clusters, but differences are also observed between stop and strident clusters even at this early stage. These results indicate that clusters are not entered in the child's lexicon as a sequence of segments but as single elements, distinct from singleton consonants.

One might hypothesize that these results were due to constraints on the output mechanism of the child, that is, his inability to produce two consonants in sequence, rather than the way in which he stores information about clusters. A study was undertaken to explore this question. Children aged three years to eight years were asked to learn, from repeated oral presentation, two sets of nonsense syllables: one containing initial clusters found in English, and one containing non-English clusters. Figure 2 indicates the most frequent sound substitutions for each of the initial clusters

English Clusters		Non-English Clusters	
KW → KL	.01	KV → KW	.001
GL → DR, GR		GZ → Z	.001
DR → GL	.001	SR → SW, STR, ST	
ST → STR	.001	DL → GL	.001
TR → ST	.01	TS → S	.001

Fig. 2. Most frequency sound substitutions for initial clusters and level of significance of frequency of occurrence of these substitutions compared to others.

in both sets and the level of significance of that response as compared to any other. As can be seen, most of the substitutions did not reflect attempts to maintain either distinctive features of speech sound segments or the segments of the original clusters. The two non-English clusters containing /s/ and /z/ in second position were recalled significantly most frequently as singleton strident clusters. No substitution for /sr/ reached even the .05 level of significance and substitutions included addition of segments. The other two clusters were made to conform to English rules. The same substitutions were found throughout the age range — of three to eight — except for /sr/. With that cluster the most frequent substitution for three-to five-year olds was /st/, for five-to six- year olds /str/ and for six- to eight-year olds /sw/. Thus, the youngest children were only observing a sequential rule of /s/ clusters whereas older children were, in addition, observing some segmental aspects, that is, that /w/ is close to /r/ in features.

These results indicate that perceptually as well as productively, the underlying representation of initial clusters is as a single consonant with some features to be used for later segmentation. This single consonant in English falls into three classes. It is a stop consonant with /w/, /r/ or /l/ features, a strident consonant with /w/ or /l/ features, and an obstruent with strident features. The particular features are realized at later stages of recall of clusters by children. This kind of processing of initial clusters also indicates that, just as in syntactic processing, chunking goes on in speech processing with later spelling out of the segments of the chunk. This basic chunk appears to be the syllable.

Research Laboratory of Electronics
Massachusetts Institute of Technology

REFERENCES

Bullowa, M., L. Jones, and T. Bever
1964 "The Development from Vocal to Verbal Behavior in Children", *Monographs of the Society for Research in Child Development* 29:101-107.

Gamkrelidze, T.B.
1966 "A Typology of Common Kartvelian", *Language* 42:69-83.

Hoffman, T.R.
1967 "Initial Clusters in English", *Quarterly Progress Report, M.I.T. Research Laboratory of Electronics* 84:263-274 (15 January).

Kornfeld, J.R.
1971 "What Initial Clusters Tell Us About a Child's Speech Code", *Quarterly Progress Report, M.I.T. Research Laboratory of Electronics* 101:218-221 (15 April).

Menyuk, P. and D.H. Klatt
1968 "Child's Production of Initial Consonant Clusters", *Quarterly Progress Report, M.I.T. Research Laboratory of Electronics* 91:205-213 (15 October).

DISCUSSION

FROMKIN (Los Angeles)
If in fact that on a certain level children seem to perceive and store clusters as single segments, I do not think one can conclude that in the adult language the phonological representation of clusters is also a single segment. Speech error data reveal that clusters such as *sk*, *st*, *br*, *gr*, etc., can be split into component segments. Thus, for the intended utterance *fish grotto*, the speaker said [frɪš gɑDo]. Or for *soup is served* the speaker said *serp* is [suwvd]. There is an abundance of such examples which reveal that at some level, the clusters must be sequences of discrete segments which can be linearly disordered in the process of speech. A model which posits these as single segments can not adequately account for the data. However, in English, affricates do not seem to behave as do other clusters — the stop and the fricative are never split in errors which suggests that affricates in this language are individual segments on an abstract level. This of course does not imply that at the articulatory stage clusters are sequences of segments. It depends on the particular stage in speech production.

MENYUK
The data obtained on children's acquisition and recall of initial consonant clusters indicate that these clusters are stored in children's lexicons as single underlying consonants with features which distinguish them from each other and singleton consonants. Later rules, that is those that come later in the acquisition of clusters and also appear to come later in the phonological recall of children, segment the cluster into its component parts. There are two possibilities to account for the error data that have been obtained with adults. The first is that, unlike children, adults store some initial clusters as sequences of matrices of features or segments and others as single underlying consonants since some clusters are split into component segments in errors and others are not. Another possibility is that all initial clusters are stored as single underlying consonants in the lexicon of adults as well as children but that segmenting rules are much more available to the adult.

Both the developmental data and the speech error occurrence data alone do not provide sufficient evidence to parcel out this question of hierarchical processing. Clearly children and adults may handle phonological processing in a different manner and simple occurrence of a certain type of error is not conclusive evidence of ordering in phonological processing. What is needed with both children and adults, and I am sure Dr. Fromkin would agree, is further experimentation to examine in detail the recall of clusters under various conditions.

FRY (London)
My only comment on Dr. Menyuk's excellent paper is to suggest that it is perhaps necessary to give rather more consideration to the interaction between the child's reception and his production of speech. I can cite one example which may make this

clear. At a certain stage, one particular child had established in reception the distinction between English *ch* and *tr* and between *j* and *dr*; he did not confuse for example *chain* with *train* or *jaw* with *draw* but in his speech production he was not yet able to produce different sounds and pronounced *ch* and *j* in all cases. Later on he learned to make *tr* and *dr* in addition and he then became for a time uncertain as to the distribution of the four phonemes and would consistently pronounce *treeks* for *cheeks*, *dreans* for *jeans* and *droe* for *Joe*. This might possibly have been noted as a defective pronunciation whereas in fact it reflected an interaction between reception and production and a consequent uncertainty as to distribution.

MENYUK
Periods of change and growth during which some old rules persist and some new rules are over-generalized to cover all instances can be found in other aspects of the child's production of language as well as speech sound acquisition. For example, when learning to expand the noun phrase the child may produce both 'I see boy' and 'I see a boys'. When applying a past tense marker to a strong verb he may produce 'goed' and 'wented'. When generating a clause construction he may produce 'I know what he doing.' and 'I know what is he doing.' At a later stage he produces structures such as 'I see the boys.', 'went' and 'I know what he's doing.' At the same time as he is producing the former types of utterances (at approximately one-and-a-half to two-and-a-half years) he will spontaneously correct some of these types of utterances when he is asked to repeat them. These corrections appear to reflect how he perceives these structures and not simply how he produces them. I believe that it is important to add that none of the developmental stages in language acquisition can be described as defective language production or pronunciation but, rather, as steps in the conclusions the child has reached about the structure of his language system. These conclusions, perceptual and productive, change as he matures.