ELECTROPALATOGRAPHY OF CONVERSATIONAL SPEECH

L. Shockey

University of Reading, U.K.

ABSTRACT

Electropalatography was used to sample natural conversational English. A tabulation was made of cases where alveolar obstruents could occur and of how these underlying consonants were realised. The results reflect large scale reduction of alveolars in conversational speech, some of which (e.g. reduced lateral contact) seem to be common to all members of the set and some of which are more particular to the class of speech sounds involved (laterals, nasals, stops, fricatives).

1. INTRODUCTION

Until recently, little or no research using electropalatography has focused on tongue-palate contact during relaxed, unselfconscious speech such as that which we use in everyday discourse. The reason for this is presumably the unease which besets phoneticians when they think about doing research on non-laboratory speech: in collecting free conversation, one cannot control for any of the variables known to influence articulation, among them segmental environment, stress, place in utterance, and word class. In addition, one never knows how many tokens of a given type will appear on any particular occasion, thereby making it hard to apply standard statistical measures to the results. Yet, surely if our goal as linguists is to model speech as it is used by ordinary people in daily life, it is vital to develop techniques for collecting and analysing data about this type of speech. Electropalatography provides an indirect but dynamic picture of articulator movement and as such is an invaluable adjunct to auditory and acoustic analysis of natural speech.

2. EXPERIMENTAL METHOD

In this study, acoustic and EPG data were collected from two subjects involved in conversation. The subjects were both longterm EPG users, having been on the team which developed the system currently in use at Reading University. They reported feeling very comfortable wearing the palate and experiencing no interference with articulation. Each of the subjects was seated comfortably in a small room and asked to talk to another member of the research team whose speech was not being monitored. The experimenter was in an adjoining room, listening to the conversation. After an initial period during which the conversants seemed to have become involved in discussion and to be producing unselfconscious output, the experimenter collected three-second samples of acoustic and EPG data. The acoustic signal was sampled at 10KHz and the EPG output at 100Hz. One minute of speech was collected from subject WJ, a West Midlands speaker with considerable Standard Southern overlay and 1.5 minutes collected from subject FG, a Standard Southern British speaker.

An impressionistic phonetic transcription of the collected corpus was done as well as a phonemic transcription. A tabulation was then made of cases of /t,d,s,z,n,l/ (the alveolar consonants involving contact in English), and each phonemic form related to both its phonetic transcription and the span of 10-millisecond EPG patterns which corresponded to it. The phonemic category provided a list of places where it would in theory be possible to find a maximally-articulated alveolar consonant; the phonetic realisations were divided into three categories: complete closure, incomplete closure, and deletion. These are very crude divisions. Complete closure was defined as the case in which every column of the palatogram indicated contact in at least one of the first four rows. Many kinds of complete closure were noted. For example, several degrees of lateral contact could be seen for everything except [l]: some showed a great deal of lateral contact, presumably indicating a high tongue position. Less side contact was visible in others, suggesting a laxer closure. The tokens with weak lateral contact were very common: this may prove to be a predictable feature of English conversational speech.

Complete closure per se cannot be said to apply to fricatives at all, since they require an incomplete closure in their production. For the same reason, the notion of incompleteness is not well-specified for fricatives: some with a very wide central channel were found, but as they were heard to produce friction, they could not be judged as incomplete.

Deletion in this case was defined as "showing no palatal contact": clearly inadequate, since a gesture of considerable proportions can be made without actually making contact with the palate.

While these categories will, therefore, have to be amended in a more detailed report, they allow us to shed some light on the behaviour of the elements investigated and so have been preserved here.

3. RESULTS

Not all underlying alveolars were fully realised, and in a pattern which was relatively similar from speaker to speaker. Table 1 shows summary data averaged over all consonants for each speaker and for both speakers combined.

TABLE 1

	WJ	%	FG	%	Both	%
all alveolars						
complete	113	71	182	69	295	69
incomplete	13	8	39	15	52	12
deleted	31	19	24	9	55	13
glottalled	3	2	20	8	23	5
total	160		265			

Characteristic realisation patterns emerged for each manner of articulation:

1. /n/ -- Reduction of /n/ can be attributed to two main factors, a) a Vn sequence is often reduced to a nasalised vowel before another alveolar consonant, and b) [n] often shows incomplete closure intervocalically.

In addition, [n] shows, in common with most of the other consonants investigated, a tendency to be articulated with a central groove before a fricative. It is a well-established tenet of phonetics that the production of the near-closure for a fricative involves finer motor control than the (theoretically) complete closures found for stops and nasals. Electropalatograms show that preparation for the groove configuration begins in preceding alveolar consonants and can sometimes be detected in vowels preceding such clusters.

2. /l/ -- In these subjects, there were two distinct realisations of /l/. One involved contact with the palate and was found syllable-initially, at the trailing end of a cluster, and intervocalically. The other involved no contact and was found at the leading end of a cluster and finally. The light or "semivocalised" closure which was noted by Hardcastle and Barry [1] in some environments was not found to be characteristic of these subjects: subject FG showed four anomalous cases, but these were a very small proportion of the total.

3. /s/ and /z/ -- These sounds tend to be preserved in some form, but (as mentioned above) often get a very wide channel in these data, implying (in agreement with the lateral contact discussed above) less raising of the tongue toward the palate than is found in citation forms.

4. /d/ -- A fully closed [d] is normally found after another non-nasal alveolar, especially word-finally when the next word begins with a vowel. The closure tends to weaken intervocalically, even if the [d] is word-final. (The resulting segment does not sound like a fricative or look like one on an acoustic display. This is presumably because there is little or no airflow through the constriction). [d] is especially prone to deletion in the environment n__C.

5 /t/ -- Fully-articulated tokens tend to be found syllable-initially, especially word-initially and especially in stressed syllables. After the alveolar nasal or fricative and intervocalically, [t] can be either fully closed or incomplete. No closure is normally found in the environment C__#C.

For both speakers, /t/ was usually realised as a glottal stop in the environment V__#C and in absolute final position.

4. DISCUSSION

Let us return briefly to the notion of a normal or target articulation. While it is clearly desirable for all speakers to be able to produce a maximally-differentiated set of alveolars in citation-form words in a laboratory situation, it seems obvious from the above that less fully realised tokens are very much a part of conversational speech and are in themselves normal. The implication for those using EPG didactically is obvious: it would be excessively demanding and in some sense even incorrect to expect maximally differentiated tokens of most alveolar consonants (in some environments) in unselfconscious speech. Variation in production which comes about not only through coarticulation with surrounding segments but also through position in the linguistic unit (syllable, word) and position with respect to stress must be taken into account. There might also be a generally lower longterm jaw/tongue setting in conversational speech, which leads to less side contact and bigger fricative grooves, and may be one of the reasons for the observed incomplete closures. (See [2] for further discussion of this question).

The latter point must be reiterated with respect to general phonetic theory: these data provide further evidence for the assertion that the physical properties of the vocal tract alone cannot account for the patterns of reduction we find in conversational speech. An /nt/# sequence behaves very differently from an /n/#/t/ sequence with respect to reduction: it is the higher-level linguistic construct which determines the possibility of phonetic variation, though the construction of the vocal tract is one of several factors which determine the nature of the variation.

BIBLIOGRAPHY

[1] Hardcastle, W.J. and Barry, W., (1989), "Articulatory and Perceptual Factors in /l/ Vocalisation in English," Journal of the International Phonetic Association 15, 3-17.

[2] Hewlett, N. and Shockey, L., "On Types of Coarticulation," in D.R. Ladd and G.J. Docherty (eds), Papers in Laboratory Phonology II, Cambridge University Press, to appear.