null
Coarticulation effects in vocal bursts are quite spectacular in speech. Such effects are most strongly indicated by the position of the lower in frequency (and generally higher in amplitude) of the two peaks of the typical vocal spectrum. An example of this phenomenon is displayed in Figure 3. Male/female differences in bursts again tend to reflect male/female differences in formants. It should be noted that the sexes are not the only possible transition of the constrictions. Differences to be noted are far more pronounced in the visual than in the auditory transition. Generally, however, these differences are produced by the clear form of transitions between the source of the speech and the pinching of the vocal folds.

A PHONIC RECOGNITION ALGORITHM

A method for automatically classifying the plosive consonants was developed from this study. In particular it was developed from the results for three speakers and tested on the remaining two. This algorithm is described in detail in [11]. The rules are speaker-independent within each sex group and prominent features are the likelihood that any unknown plosive was produced in any of the three primary places of articulation. The rules primarily involved three fundamental measurements - measures of burst mass, measures of burst mass that transition amplitude, and measures of burst transition amplitudes. The rules are constructed such that a variety of speaker variation effects (such as burst non-production) are allowed for. Using these rules it was found that 92% of the plosive consonants tested were correctly and uniquely classified. 

BURST PHENOMENA

It was found that the burst spectral shapes fell into the three categories. Labial bursts were diffuse and of low energy. Alveolar bursts were generally not so diffuse and were higher in energy. Alveolar bursts are more prominent in the 0.7-2.0 kHz region while labial bursts tend to be in the 1.0-3.0 kHz region. Labial bursts characteristic display two, narrow-bandwidth peaks. The more prominent of these occurs in the 0.7-2.0 kHz region and the slightly smaller one occurs in the range 3.0-5.0 kHz. Alveolar bursts characteristic display two, narrow-bandwidth peaks. The more prominent of these occurs in the 0.7-2.0 kHz region and the slightly smaller one occurs in the range 3.0-5.0 kHz. Alveolar bursts characteristic display two, narrow-bandwidth peaks. The more prominent of these occurs in the 0.7-2.0 kHz region and the slightly smaller one occurs in the range 3.0-5.0 kHz. Alveolar bursts characteristic display two, narrow-bandwidth peaks. The more prominent of these occurs in the 0.7-2.0 kHz region and the slightly smaller one occurs in the range 3.0-5.0 kHz.

GENERALIZING THE RULES

The plosive classification rules described above were incoherent in 15% of recognitions, i.e. the decision made along with a plosive classification algorithm would depend primarily on burst location and to a lesser extent on burst mass. It was found that the overall recognition performance in the general situation and that many of the reference rules (such as those allowing for burst non-production) were correct. The categories of coarticulation that might have been classified by these rules were not, however, equal in terms of success. For example, as the one described here is limited unless preconceived coarticulation algorithms are good.

As an example, 25% of all plosive productions in continuous speech were classified as 'correct' in 30% of recognitions. This situation is only of limited importance. Nevertheless they have proved a useful guide for other contexts and results for the classification of plosive consonants in step consonants. The results of this study are consistent with the finding of other researchers who have used similar techniques.

Figure 2: Typical examples of bilabial, alveolar and labiodental consonants in four vowel contexts.

REFERENCES


Figure 3: Plosive consonants showing coarticulation effects for a male speaker.

Figure 4: lcerele showing the distribution of the plosive consonants for the two men. Fv points that differ in all cases and hence they represent the TFL range.