Abstract

An acoustic property that distinguishes velar consonants from labial, alveolar, and dental consonants is a prominent midfrequency "compact" spectral peak, usually in the frequency range of 800-4000 Hz. In a series of perceptual experiments, synthetic syllables with initial voiced and voiceless stop consonants were generated, and the spectral characteristics of the consonant burst were systematically manipulated to yield various degrees of prominence of the mid-frequency spectral peak. From listener responses to these stimuli, we have determined that the property of compactness depends in part on the amplitude of the prominent spectral peak in relation to a peak at about the same frequency in the following vowel. Spectral analyses of a number of naturally spoken stop consonants in English have shown that the amplitude characteristics of the mid-frequency spectral prominence of the burst are consistent with the perceptual data. However, the degree of prominence often shows fluctuations throughout the region encompassed by the burst and voicing onset in the following vowel.

1. Introduction

The most distinctive acoustic characteristic of velar stops is usually said to be a compact spectral prominence, in the mid-frequency range of 800-4000 Hz. As a result we see the smoothness of the burst and vowel onset of a naturally-spoken /g/ sound, together with the waveform. The burst spectrum has the classical compact mid-frequency prominence. Another attribute of the pattern in Fig. 1 is that the amplitude of the spectral peak in the burst is comparable (within about 5 dB) to the amplitude of the corresponding spectral peak in the vowel. This characteristic of the burst in relation to the vowel is consistent with data reported by several investigators [1, 2]. Velar stops also have a number of secondary characteristics, such as bursts that are longer, and first-formant transitions that tend to be slower than those for bilabials and alveolars. Nevertheless, spectra of velar stops vary a great deal, and the concept of "compactness" is poorly understood.

This paper describes some preliminary work in a planned series of studies of the acoustic characteristics of velar stops. We asked two questions: both of which focus on spectral rather than on temporal properties. First, can we synthesize an acceptable velar stop simply by manipulating the spectrum of the burst alone, and if so, what are the critical acoustic characteristics of such bursts? And second, to what extent are compact characteristics observable in naturally-spoken syllables? The focus of interest is the release burst and the first few milliseconds of the following vowel in syllable-initial stops.

2. Perceptual experiment

The stimuli for a perceptual experiment consisted of a series of synthetic consonant-vowel syllables. We constructed acoustic continua of bursts such that, when these bursts are followed by minimal vowels, we heard velar stops at one end of the continuum, and either alveolars or bilabials at the other end, depending on the continuum. The various bursts were synthesized manipulating the amplitudes of noise-excited formants in parallel synthesis.

Figure 2 shows short-time smoothed spectra of bursts at the extremes of the two continua—the velar-bilabial in the upper panel, and the velar-alveolar in the lower panel. If the velar-alveolar set of bursts, the classically compact shape of the spectrum labeled /g/ contrasts with the diffuse, rising spectrum, /d/, that is typical of alveolars. The variations in spectrum shape were achieved by changing the amplitude of excitation of F2. For the velar-bilabial set, formant amplitudes were altered so that the compact /g/ spectrum was made flatter and slightly falling, as for a bilabial.

Figure 3 shows the results for the CV stimuli of the velar-bilabial continuum and the velar-alveolar continuum for nine subjects. Forced-choice categorization functions for each continuum had a reasonably sharp crossover between 100% velar and 100% bilabial or alveolar responses, indicating that most listeners could classify sounds in terms of place of articulation using only the burst spectrum, perhaps relative to certain characteristics of the vowel spectrum. There are some differences between the responses for the velar-alveolar continua with and without aspiration, presumably a consequence of the closer proximity of the vowel onset to the burst for the voiced continuum.
time, followed by a vowel with suitable transitions. Consequently, we have reexamined the acoustic properties of spoken velar consonants, particularly the fine structure of the short-time spectrum through the burst and into the onset of voicing.

### 3. Acoustic analysis of natural speech

We have looked at CV syllables spoken by several talkers of British or American English saying /gi/, /ge/, /ga/, and /gə/ (as in ‘gala’). Several types of spectra were made of the burst and at least the first two periods of the vowel, including Fourier transforms, LPC spectra, and the output of certain auditory models. Spectra were made in successive 5 ms steps and additionally, for some syllables, in smaller steps. Many of the spectra conform closely to the classical [compact] description for the burst. But as other investigators have also found [2], a substantial minority deviates from the classical picture. However, almost all of these so-called deviant or atypical utterances, have compact properties during at least some part of the burst or vocalic onset. Two of the most common types will be shown here.

**Figure 5** shows LPC spectra at successive 5 ms intervals from the beginning of a second syllable, with a speaker from the same talker. This sequence is an example of two formants being close together at the beginning of the transition, followed by a vowel with suitable transitions. Consequently, we have reexamined the acoustic properties of spoken velar consonants, particularly the fine structure of the short-time spectrum through the burst and into the onset of voicing.

### 4. Conclusions

In summary, we have seen in the perception experiments that velar stops in a CV syllable with steady-state formants are heard if the burst has a midfrequency spectral prominence with an amplitude at least as great as that of the corresponding peak at vowel onset. The analyses of natural speech show that the compact prominence is typically present in the burst spectrum, or it may be only intermittent, or it may be more evident in the vowel onset than in the burst. These data suggest that compactness should be defined in terms of the prominence of a peak in the average burst spectrum, rather than in terms of the occurrence of prominence in the short-time spectrum in at least some region of the syllable onset, whether it be in the burst or in the onset of voicing. One possibility is that the perception of velar in the consonantal release is enhanced if there are regions in the release phase where a compact spectral prominence is embedded in a context that has reduced compactness or prominence.

If these preliminary observations are confirmed on a larger dataset, then the next task is to begin to describe compactness more precisely through further perception experiments. If we can express compactness in terms of the amplitude, bandwidth, frequency range and time-course of a midfrequency peak relative to adjacent spectra, then we may be on the way to coming up with a description that subsumes burst and transition information under one umbrella.

### 5. Acknowledgements

This work was supported in part by grants from the National Institute of Neurological and Communicative Disorders and Stroke and from the National Science Foundation.

### 6. References