Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 409–411
(S. Karger, Basel/New York 1965).

Speech Transmission Laboratory Royal Institute of Technology (KTH), Stockholm

Information Bearing Aspects of Formant Amplitude

By JÁNOS MÁRTONY and GUNNAR FANT, Stockholm

A formant may be specified by its frequency, amplitude, and bandwidth. Most of the experimental studies of formant structure up-to-date have been concerned with formant frequencies. In recent years communication engineers have shown a rising interest in studying formant bandwidth. This is motivated by the role of the frequency and the bandwidth as the measurable counterparts of the imaginary and the real part of a complex number, the pole entering the theory of resonance phenomena. Formant amplitudes, however, have not been studied to the same extent. Data on formant amplitudes of English vowels have been given by *Peterson* and *Barney*[1] and corresponding data on Swedish vowels by *Fant*[2].

Formant amplitudes are related to formant bandwidth to the extent that an increase in formant bandwidth, everything else held constant, results in a reduction of the formant amplitude. Formant amplitudes are in addition dependent of the overall pattern of formant frequencies and possible antiresonances (antiformants) so that a formant amplitude increases, when it is shifted to a position closer to an adjacent formant and it decreases as it is made to approach the frequency location of an antiresonance. A shift up in the frequency of the first formant is associated with an increase in the intensity level of the entire spectrum above the first formant including all formant amplitudes. This is the natural mechanism whereby a more open (high F_1) sound is more intense than a more closed (low F_1) vowel phonated with the same vocal source effort.

In addition to these changes of formant amplitudes which are a function of the articulation, i.e. the vocal cavity configuration, and thus are fully predictable the main source of variations is the phonation, i.e. the particular spectrum of the raw material supplied by the vocal cord source or by any other primary source. The source characteristics changes with the speaker, his voice effort, pitch level, and the prosodic stress pattern within connected speech. They are to some extent also dependent on the articulation. Spontaneous variations in the glottal waveform and thus in the corresponding source spectrum occur frequently. Recent studies[3] have revealed considerable fluctuations of formant amplitude in connected speech which are related to the particular ratio of voice fundamental frequency F_0 to a formant frequency F_n, particularly F_1. These fluctuations involve nonlinear interaction between the source and the filter.

It is well known that a vowel retains its phonetic identity irrespective of formant amplitude changes providing these are within reasonable limits (10–20 dB) and the pattern of formant frequencies remains constant. Individual variations typical of a speaker's voice category may be studied.

In connected speech the transition from a vowel to a consonant is typically accompanied by a change in formant amplitudes. In case of a transition from a nasal consonant to a vowel the instant of oral opening is associated with a radical change of formant amplitudes. Typically the second formant is much weaker in the consonant than in the vowel. This is one of the cues of nasal consonants.

The object of the present study was to see if appropriate formant amplitude changes in synthetic speech introduced via formant bandwidths could create the effect of a transition from a nasal consonant to a vowel.

From our experience of speech analysis and synthesis and earlier investigations, e.g. *Fant*[4], *Fujimura*[5], it would be expected that a correct reproduction of formant amplitudes would enable a listener to perceive a synthetic syllable as lateral + vowel if the bandwidths were low and nasal + vowel if the bandwidths were larger, i.e. when formant amplitudes were accordingly decreased. A clear tendency in this direction was observed.

Our study was devoted to the nasals [m] and [n] in CV syllables with the vowels [i] [u] and [a]. The synthesis was made on OVE II, a synthesizer with the formant circuits in cascade. In addition to the normal control via conducting ink lines drawn on a function generator there were relay controls incorporated for introducing changes in formant bandwidths and thus in formant levels.

Our preliminary results may be summarized as follows.

[ni] A low level of the second formant in the nasal murmur is the primary cue. The nasalization cues in the form of bandwidth increase of the higher formants were not significant.

[na] The correct reproduction of formant amplitudes in the nasal murmur was not sufficient. The nasalization cue in the vowel in the form of a decreased first formant amplitude was necessary. Without this modification the syllable is heard as [da] or [la] and the perception of [l] is further enhanced by reducing the bandwidth of the second formant within the consonant interval.

[mi] The low L_2 in the nasal segment is important. If this level is raised the nasal shifts to [l] or [b]. The nasalization of the [i] was not important as a cue for the nasal consonant.

[ma] Both the nasal murmur and the vowel are of importance for the identification of the nasal.

[mu] The optimum synthesis required a low L_2 in both the nasal
[nu] and the vowel. The [mu] was, however, not quite natural.

The increase of formant bandwidth as a means of reducing formant amplitude has the unavoidable effect of spectrum flattening which does not allow the reproduction of a valley in the 700 c/s range of the nasal murmur.

References

1. *Peterson, G. E.* and *Barney, H. L.:* Control methods used in a study of the vowels. J. Acoust. Soc. Amer. *24:* 175–184 (1952).
2. *Fant, G.:* Acoustic analysis and synthesis of speech with applications to Swedish. Ericsson Technics *15,* 1: 3–108 (1959).
3. *Fant, G.; Fintoft, K.; Liljencrants, J.; Lindblom, B.* and *Mártony, J.:* Formant amplitude measurements. To appear in J. acoust. Soc. Amer. (November 1963).
4. *Fant, G.:* Acoustic theory of speech production (Mouton, 's-Gravenhage 1960).
5. *Fujimura, O.:* Analysis of nasal consonants. J. acoust. Soc. Amer. *34:* 1865–1874 (1962).

Authors' address: Dr. J. Mártony and Dr. G. Fant, Speech Transmission Laboratory, Royal Institute of Technology, *Stockholm 70* (Sweden).