Dept. for Speech, Music and Hearing

Quarterly Progress and Status Report

Pole-zero matching techniques

Fant, G. and Mártony, J.

journal: STL-QPSR
volume: 1
number: 1
year: 1960
pages: 014-016

KTH Computer Science and Communication

http://www.speech.kth.se/qpsr

F. POLE-ZERO MATCHING TECHNIQUES

Our work in this field is still in an exploratory phase. Vowels are studied by means of the anti-resonance filter techniques mentioned in section I A. Most of the pole-zero matching of fricatives was made on a graphical basis comparing the spectra of human samples with spectra synthesized numerically from tabulated data of elementary pole and zero curves. Analog methods based on the use of networks with variable poles and zeros are also being investigated. A standard circuit for representing a pole-zero pair has recently been developed (Kringlebotn). It is based on the continuous variation of an inductance by means of feedback amplifier techniques.

A match of fricatives in terms of two poles and one zero is generally sufficient for retaining a high standard of speech quality in a formant-coded synthesis (OVE II).

A matching of the fricatives [s] and [f] in terms of two poles and one zero is shown in Fig. I-7. The measured samples pertain to sustained sounds analyzed by a closed loop process with a wave analyzer of 125-c/s bandwidth. The pole at 2700 c/s and the zero at 2500 c/s of the fricative [f] constitute a bound pair with but small contribution to the spectrum. It is of interest to see that the spectrum level rises all the way up to 12 kc/s which was the upper limit of analysis. Spectra of [f] vary much owing to the particular coarticulation and to the degree of labiodental constriction.

The main peak of the [s]-spectrum of Fig. I-7 is associated with the pole at 5800 c/s. The second pole at 8000 c/s contributes to build up a proper spectrum level at higher frequencies. The zero at 4500 c/s is placed higher than the corresponding zero in the measured spectrum in order to preserve a correct level ratio between the main formant and the low frequency part of the spectrum.

An additional inventory of two pole-zero pairs were added for matching the [s]-spectrum of Fig. I-8. One of these bound poles, the one at 2500 c/s, corresponds to F3 and the one at 4200 c/s to F5. These weaker formants do not appear to be necessary for the synthesis of a good [s].

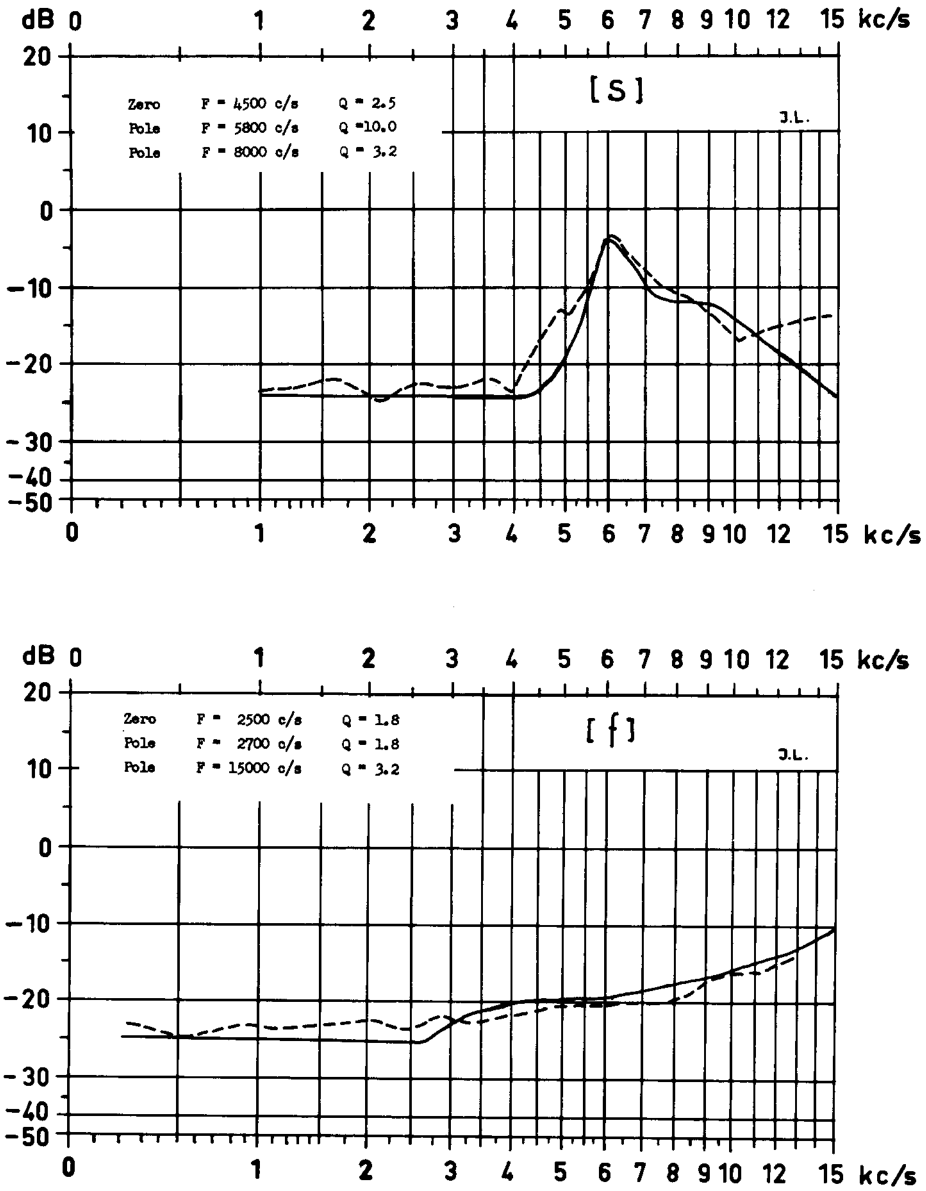
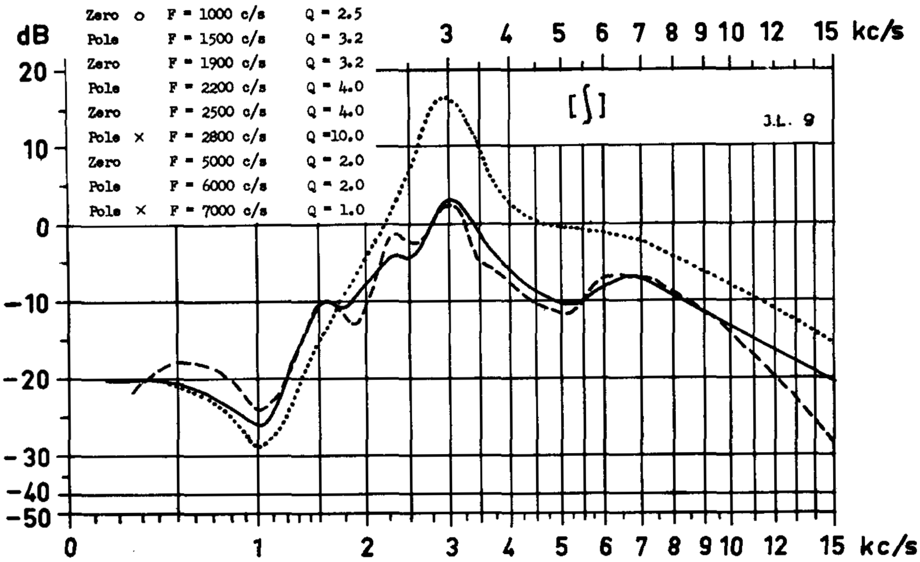


Fig. I-7 Measured spectra (broken lines) and two-pole-one-zero synthetic approximations (solid lines) of the fricatives [s] and [ʃ].



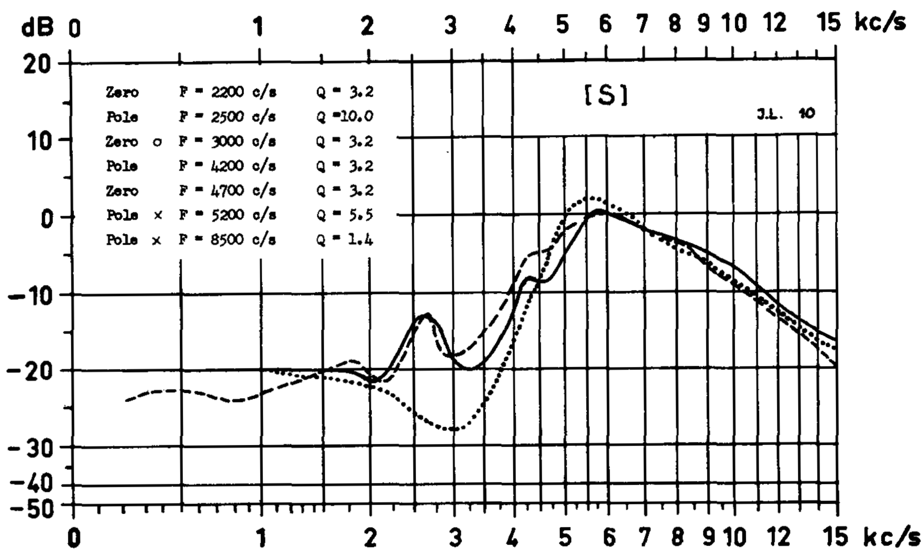


Fig. I-8 Pole-zero matching of [ʃ] and [s]. Measured spectra are represented by broken lines, two-pole-one-zero approximations by dotted lines and more complete synthetic representations comprising additional bound pole-zero pairs by solid line curves. Free poles are marked × and free zeros are marked o.

The spectrum of a fricative [ʃ] and its pole-zero approximation is demonstrated in the lower part of Fig. I-8. The essential feature of this particular palatal retroflex sound is a free zero at 1000 c/s and a free pole at 2800 c/s and one at 7000 c/s. A detailed match employing three additional pole-zero pairs associated with the relatively suppressed F2, F3, and F6 provides a match within a few dB from 300 c/s to 12 kc/s. The dotted curve on the figure pertains to the approximation in terms of the two free poles and the free zero alone. It is apparent that the resulting exaggeration of the relative level of the main peak is due to the absence of the high-frequency attenuation inherent in the two bound pole-zero pairs. This effect has been predicted in earlier theoretical work.(1) In agreement with results from those earlier studies (1) it is apparent that the synthesis can be made on the basis of a relatively flat source spectrum.

The pole-zero matchings performed in Fig. I-7 and I-8 allow a simplified structural comparison of the sounds [f], [s], and [ʃ]. There is a similarity between [s] and [ʃ] in so far as the spectra of both possess a free zero of a frequency lower than that of the two free poles. This free zero contributes effectively to the high-pass structure of the spectrum above the zero the significant part of which extends approximately 2000 c/s lower down in frequency for [ʃ] than for [s]. The mode spectrum of [f] does not possess a free zero and the only free pole is located at very high frequencies and is generally heavily damped. This pattern explains in part the relatively low overall intensity of [f].

An alternative interpretation applicable to the theory of distinctive features (2) would be to oppose [ʃ] to [s] and [f] as being the only sound that has a free zero below or close to F_2. This is a requirement for emphasizing formants F3 and F4 and also F2 if the zero is well below F_2 and thus a formant area in the consonant not far above the mean pitch of the upper formants of a following vowel. This conforms with the criterion of a major energy concentration in a centrally located peak as required by the definition of compactness. After correction of the [s]- and [f]-spectra for the relatively low sensitivity of the ear in the high-frequency region it is apparent that the [s]-spectrum has a higher center of gravity than the [f]-spectrum and [f] is thus grave compared with [s]. However, in some languages (e.g., in Swedish) it is feasible to oppose [ʃ] to [s] as being more flat (shift down of the spectrum) while other fricatives, e.g., [ɕ] take the place of the compact member of the system.

G. Fant, J. Mártony

(1) Fant, G.: Acoustic Theory of Speech Production ('s-Gravenhage, 1960).

(2) Jakobson, R., Fant, G., Halle, M.: "Preliminaries to speech analysis. The distinctive features and their correlates", M.I.T., Acoustics Laboratory, Techn. Rep. No. 13 (1952); 3rd printing.