CANADIAN ACOUSTICS/
ACOUSTIQUE CANADIENNE
1987, 15 (2): 5 - 13

APPLICATION OF MONAURAL FUSION TECHNIQUE FOR EXPLORATION OF POTENTIAL INTERACTION BETWEEN CHANNELS OF PHONETIC ANALYSIS

Dale C. Stevenson*, Anton J. Rozsypal and John T. Hogan

Department of Linguistics, The University of Alberta
Edmonton, Alberta, Canada T6G 2E7

ABSTRACT

Interaction of detectors in speech recognition has been investigated using a monaural superposition technique which permits simultaneous probing of individual detectors. Waveforms of two natural speech tokens of similar temporal patterns are mixed after careful period-by-period alignment. A stimulus continuum is generated by varying the relative amplitude of the two components. In monaural presentation certain speech stimulus types fuse in perception so that only one of the two components of the mixed stimulus is perceptible. Identification experiments have been conducted on the /bæ/-/dæ/, /ra/-/la/, and /i/-/æ/ continua. For all tested consonant pairs monaural fusion takes place. In the case of the vowels tested, fusion has not been observed. Instead, both component vowels are perceived simultaneously even for extreme level differences of 26 dB. This suggests that detectors responsible for recognition of vowels are essentially independent, while detectors for stop consonants interact in an inhibitory manner.

SOMMAIRE

L'interaction entre des détecteurs de traits phonétiques a été investiguée à l'aide d'une technique de superposition monaurale de paires de syllabes de type CV ou V produites par une voix humaine. La durée des syllabes dans chaque paire est semblable. Après un alignement des périodes correspondantes dans les deux syllabes, les deux signaux sont additionnés. Un continu de stimuli est créé en variant l'amplitude relative des deux composantes. Quand certains des stimuli composés sont présentés de façon monaurale, seulement une des deux composantes est perceptible. Les continus /bæ/-/dæ/, /ra/-/la/ et /i/-/æ/ ont été utilisés dans des tests d'identification. Dans le cas des stimuli de type CV les auditeurs ont eu l'impression d'un stimulus simple, tandis que dans le cas des stimulus de type V les deux composantes restaient perceptibles, même quand la différence de niveau était de 26 dB. Ceci suggère que les détecteurs de voyelles agissent de manière indépendante, tandis que l'excitation du détecteur d'une consonne réduit le niveau d'excitation des détecteurs d'autres consonnes.

1. INTRODUCTION

From the results of various experimental paradigms used for investigation of categorical perception, such as storage of speech stimuli in auditory and phonetic memory, masking, dichotic listening, and reaction time in identification and discrimination, it appears that vowels are perceived differently from consonants. An extensive review of this topic can be found in Studdert-Kennedy (1976). These differences seem to be related to the differences in temporal and spectral structures of consonant and vowel stimuli and to differences in the perceptual processing of these types of stimuli. Similarly, as will be seen in the following experiments, a monaural superposition technique, described below, shows fundamental perceptual differences between selected consonants and vowels.

It has been proposed in speech perception studies that speech recognition in channels of phonetic analysis is mediated by acoustic or phonetic feature detectors. The major support for this idea has come from selective adaptation studies. Shift of the category boundary toward the adapting stimulus on a two-category continuum can be explained as a result of desensitization of one of the two detectors which span the continuum (Eimas and

* This paper is based on parts of the Ph.D. dissertation submitted by D.C. Stevenson to the Faculty of Graduate Studies and Research of the University of Alberta. His present address: MacDonald, Dettwiler and Associates, Vancouver, B.C., Canada.

Corbit, 1973; Miller, 1975, 1977; Ainsworth, 1977). Such a two-detector model has been quantitatively formulated by Elman (1979) and its properties further analyzed by Rozsypal, Stevenson, and Hogan (1985).

For the experimental hypothesis, it has been assumed that the perceptual differences between consonants and vowels are due to different organization of the system of channels of analysis. Speech signals for which monaural fusion takes place are presumed to be processed by channels interacting in an inhibitory way. That is to say, when the excitation of one of the two channels only slightly exceeds that of the other channel, then only the stronger stimulus is heard. Hence it is presumed that the channel more strongly excited inhibits the response of the other channel. On the other hand, channels of analysis dedicated to detection of non-fusing stimuli can be regarded as essentially independent.

To allow testing of this hypothesis, a monaural superposition technique was developed to determine the degree of interaction between posited channels of phonetic analysis for different speech types. To make a distinction between the technique and the resulting perceptual effect, the technique will be referred to as "monaural superposition." The term "monaural fusion" will be reserved for the perceptual phenomenon.

The technique is a modification of the dichotic fusion technique used by Halwes (1969), Ades (1974), Cutting (1976), and Repp (1976, 1980). For the monaural superposition technique, the stimuli are generated by mixing, in varying amplitude ratios, two or more tokens from distinct phonetic categories. Presentation is monaural. Waveforms of the mixed stimulus components must have a similar temporal structure so that they can be exactly aligned in time. In a two-category identification task, the composite stimulus is a mixture of an amplitude weighting α for one speech signal plus a weighting $(1-\alpha)$ for the other. The weighting parameter α defines a continuum which ranges from $\alpha=0$, representing the original token from one speech category, to $\alpha=1$, representing the other.

The resulting waveform contains the phonetic cues for both of the component speech categories. Each of the detectors involved is simultaneously stimulated by its respective typical speech input. In the experiments with consonant stimuli reported here, when the components are combined with approximately equal weight, the resultant signal is found to be perceptually ambivalent, and is perceived as either one or the other stimulus component.

2. EXPERIMENTAL METHOD

2.1 Subjects

Five volunteers, faculty members and graduate students with some phonetic training, including two of the authors, served as subjects in this study. Audiometric examination of the subjects revealed no marked hearing loss in the relevant frequency range below 4 kHz.

2.2 Generation of the test stimuli

All signal processing and subject testing were carried out with a special-purpose operating system designed for the DEC PDP-12 laboratory computer (Stevenson and Stephens, 1978). The quality of the resulting stimulus depends on the precision of the alignment. Misalignment of individual glottal periods may result in a detectable low-frequency noise in the mixed stimulus. The temporal alignment procedure will be described here briefly with the stimulus pair /bæ/-/dæ/ as an example. For more details, including spectrograms of the mixed stimuli, the reader is referred to Stevenson, Hogan, and Rozsypal (1985).

Stimuli /bæ/-/dæ/ have similar temporal and formant structure. Overall durations of both stimuli and the durations of initial formant transitions are comparable. When the same following vowel is chosen, both sets of formant trajectories asymptote toward the same target frequencies.. Both stimuli have a rising F1 onset transient. They are distinguished primarily by the initial frequency and direction of F2 and F3 transitions both of which rise for /bæ/ and fall for /dæ/. In reference to different segments of the stimuli, the following notation is adopted: /bæ/ stands for the entire CV syllable, /b/ for the initial nine periods containing the formant transitions extracted from that syllable, and /æ/ for the tenth through the last pitch period of the CV waveform comprising the steady-state vowel.

The feasibility of aligning the stimulus components depends in the first place on obtaining tokens of the two test stimuli with sufficiently close frequency and time parameters. A male native Canadian English speaker recorded multiple tokens of all of the stimulus components with fundamental frequencies of about 100 Hz and approximately equal steady-state vowel formant values. One pair, /bæ/ and /dæ/ for the first experiment, was then selected with the best matching fundamental frequencies and steady state formant values determined from spectrograms and waveform plots. The chosen tokens were digitized. The first nine glottal periods of these two waveforms were extracted and stored separately. The /b/ waveform with the higher fundamental frequency was arbitrarily selected as the "standard" segment for the period alignment. The signals /b/ and /d/ were segmented into individual pitch periods at zero crossings preceding the major peak. Each /d/ period was then cross-correlated with its corresponding /b/ period in order to determine the temporal shift for best match. The trailing ends of the aligned /d/ periods were then truncated to match the length of their /b/ counterparts. A new /d/ formant transition signal was created by concatenating these time-shifted and truncated /d/ periods. Thus the pitch contours of the resulting /d/ waveform and the original /b/ waveform optimally matched on a period-for-period basis. Formant frequencies are not affected by this signal operation. The new /d/ was then scaled in amplitude so that its computed overall intensity was equal to that for the /b/.

These /b/ and /d/ signals, together with the steady-state vowel /æ/ extracted from /bæ/, were used as the basic components for construction of the test stimuli. The two formant transitions were added together point by point with a relative amplitude weighting α/b/+(1-α)/d/ and then concatenated to the steady-state vowel nucleus /æ/. The linear weighting of the /b/ and /d/, along with the convergence of both of these waveforms to the same steady-state vowel ensured continuity of the amplitude envelope of the mixed formant transitions and the vowel nucleus.

Exactly the same procedure was followed for the construction of the /ra/-/la/ stimuli. In the case of the /i/-/æ/ vowel stimuli, the full length of one vowel signal was matched period-by-period to the other vowel signal in identical manner used for aligning of segments containing the formant transitions of the consonantal stimuli.

2.3 Apparatus

The stimuli were recorded in an acoustically isolated chamber using a Sennheiser MD 421N microphone and a TEAC AR-70 tape recorder. The same tape recorder was used also for subsequent digitization of the audio recordings by a DEC PDP-12 computer. A Rockland Series 1520 Programmable Dual Hi/Lo Filter with 24 dB/oct slope was used as an anti-aliasing filter, set to the Butterworth characteristics and pass-band from 70 to 7000 Hz. The sampling rate was 16 kHz and the resolution of the A/D and D/A converters was 10 bits. For technical considerations of digital processing of speech signals see Rozsypal (1976).

The stimuli were delivered to a remote listening station over lines with a signal-to-noise ratio better than 55 dB and then smoothed by the same filter identically set as for anti-aliasing. The output of the filter was amplified by a Braun CSV 250 power amplifier modified for headphone output. Its output was fed into a bus which serviced Telephonics TDH-49 headphones in MX-41/AR cushions. The frequency response of the complete audio channel was 160-4000 Hz $\pm$1 dB (100-5000 Hz $\pm$2 dB). The listening level was set at 80 dB SPL for a continuous 1000 Hz sine wave with an RMS voltage equal to the RMS voltage for the steady-state vowel /æ/. The calibration of the apparatus was aided by the Brüel & Kjær Beat Frequency Oscillator 1022, Artificial Ear 4153, Pistonphone 4230, and Level Recorder 2307. The listening level was calibrated prior to each testing session.

The procedure for generating the stimuli was as follows: The pair of formant transitions of the two signals being combined were amplitude scaled by factors of α and (1-α), respectively. The value of α for each presentation was read in from a file of 21 randomized numbers representing equal steps of $\Delta\alpha$=0.05 from 0 to 1. The scaled formant transitions were added together point-by-point and concatenated to the steady-state vowel.

2.4 Testing procedure

The stimuli were presented on-line, monaurally to the right ear, to one or more subjects in a sound-treated room. The time required for the loading, scaling, addition and concatenation of the stimulus segments was approximately 1.5 s. After the stimulus was played back, response switches at the remote listening stations were monitored until all subjects responded. The response was a choice from two alternatives: "bæ" and "dæ" for the identification of stops, "ra" and "la" for liquids, and "i" and "æ" for vowels. The program then proceeded with the next presentation. The interstimulus interval was thus somewhat variable, averaging about four seconds.

The presentation was divided into blocks of 25 stimuli. All 21 conditions were exhausted before a new randomization cycle was started. A one-second 1000 Hz tone was played back at the end of each block followed by five seconds of silence. In a test run, each one of the α conditions was presented ten times for a total of 210 stimuli. The first block contained 15 practice stimuli. A testing session lasted about 20 minutes. Each subject participated in three testing sessions, held only once in a given day.

3. IDENTIFICATION RESULTS

A series of experiments was carried out to determine whether monaural fusion takes place for the following three speech categories: stops, liquids, and vowels. Except for the vowel identification experiment, the presentation paradigm in the following experiments was the same.

3.1 Identification of stops /bæ/-/dæ/

The first experiment involved identifications of the /bæ/-/dæ/ stimulus pair. The endpoint stimuli, being natural speech tokens of /bæ/ and /dæ/, always resulted in 100% correct identifications. Throughout the rest of the continuum the stimuli were perceived by all five subjects as clear instances of either /bæ/ or /dæ/. Each of the five subjects was tested three times. Identification results for the five subjects are shown in Figure 1, where the "bæ" identification rate is plotted as a function of an amplitude ratio of the /b/ component to the /d/ component expressed in decibels by the formula $20\log[\alpha/(1-\alpha)]$. None of the identification trials with stop stimuli presented any difficulty for the subjects. They heard only one clear representation of one of the stimulus components. No phonetic intrusions were reported in post-test interviews.

3.2 Identification of liquids /ra/-/la/

A stimulus continuum constructed from tokens of /ra/ and /la/ was presented to the subjects. The results are plotted in Figure 2. Although the identification functions appear as steep as in the stop consonant experiment, the subjective impression in this case was different. Whereas in the stop experiments only one of the two competing stimuli was consistently evident, for the /ra/-/la/ experiment there were cases observed where one stronger stimulus was perceived simultaneously with the other stimulus of lower loudness in the background. In post-test interviews, two of the subjects reported hearing on very few occasions additional perceptual blends: one heard either "bra" or "bla" and the other "bla".

3.3 Identification of vowels /i/-/æ/

Following the stimulus generation procedure of the preceding experiments, vowels /i/ and /æ/ were mixed. Initially, the response paradigm used in the consonant identification experiments was also applied for the testing of vowel identification. It became immediately obvious that this paradigm was unsuitable for the testing of vowels. The stimuli just simply did not fuse perceptually and, therefore, the subjects were unable to give a single category response. They heard both stimulus components simultaneously, except for the two extreme cases of $\alpha=0$ or $\alpha=1$ representing the original single vowel stimuli. This was verified with informal up-and-down runs of the amplitude parameter α in which the subjects were required to state whether they heard one or two of the vowel stimuli. No identification curves could thus be constructed. The vowel with the higher amplitude factor determined the stronger percept, which was accompanied by a simultaneous weaker percept of the other vowel. Since the step size was $\Delta\alpha=0.05$, this means that the weaker vocalic stimulus was audible down to at least 26 dB below the stronger one.

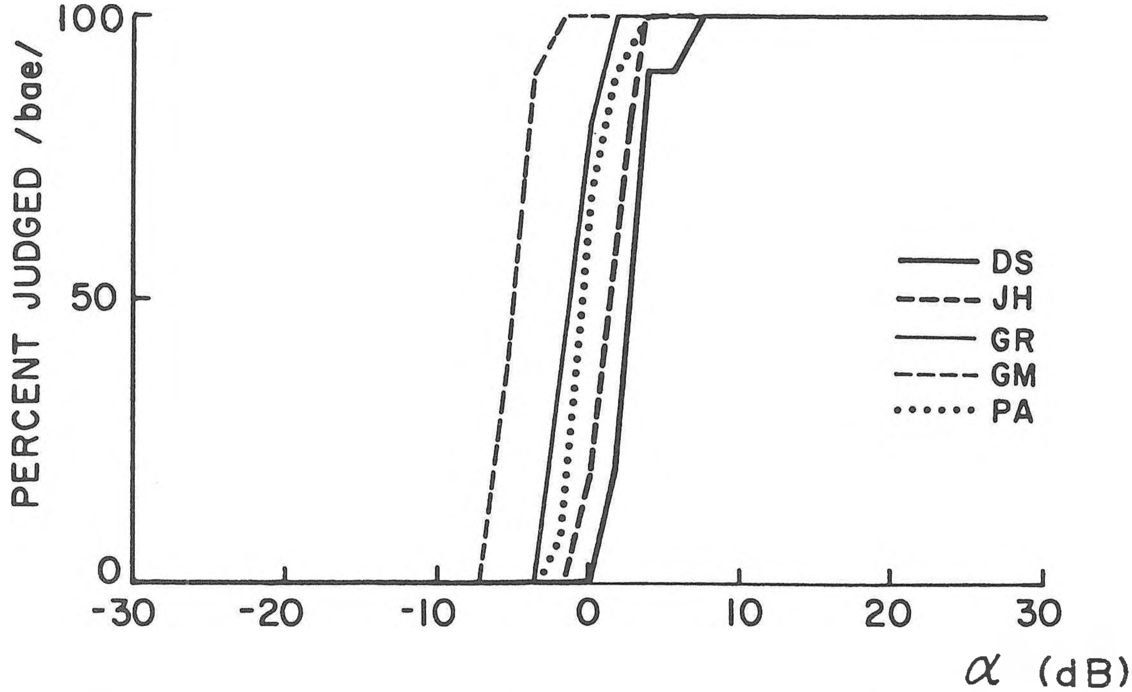


Figure 1. Identification by five subjects of the composite stimulus /bæ/-/dæ/ as a function of the amplitude ratio α of the /b/ to /d/ component expressed in decibels.

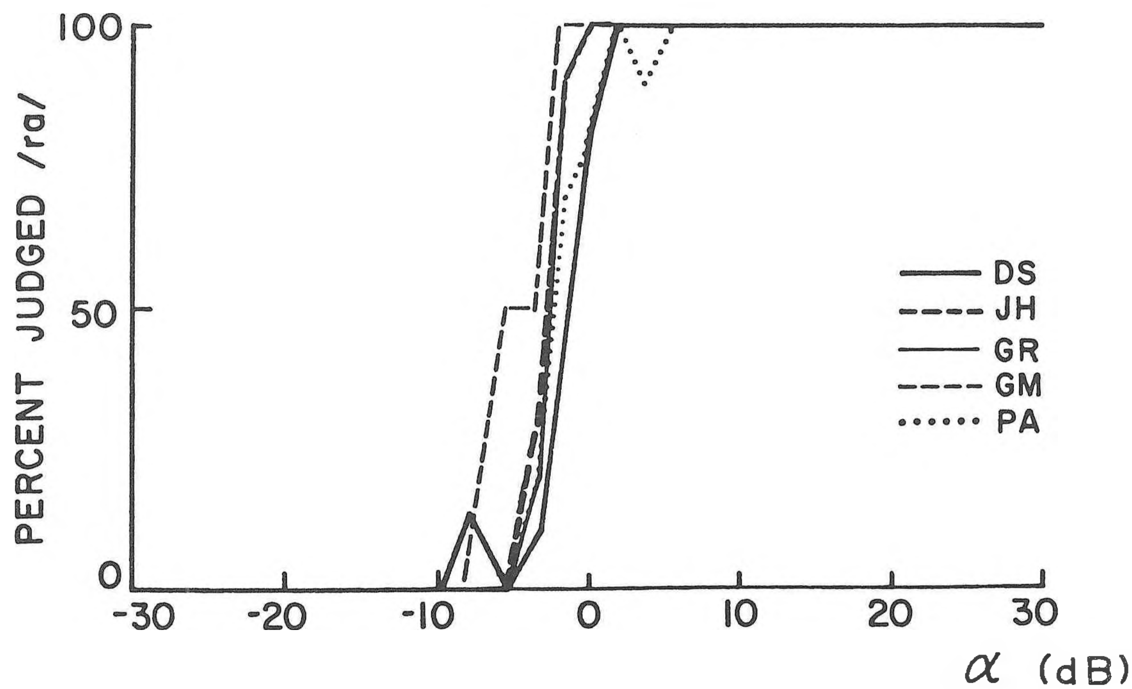


Figure 2. Identification by five subjects of the composite stimulus /ra/-/la/ as a function of the amplitude ratio α of the /r/ to /l/ component expressed in decibels.

4. DISCUSSION

The evidence from the limited number of stimulus pairs tested in this study suggests a fundamental difference between the perception of vowels and stop consonants. This difference can be viewed as being due to the organization of the decision stage of the recognition mechanism. In the case of stops, monaural fusion takes place consistently since only one of the mixed stimuli is perceived. The decision stage of the auditory analyzer selects one and only one of the alternatives. Thus it appears that the channels of phonetic analysis inhibit each other so that only the channel with the stronger stimulation is selected. On the other hand, stimuli such as vowels, for which monaural fusion is not observed, can be hypothesized as being processed by separate and essentially independent auditory channels. Each vowel component of the composite stimulus is independently recognized provided it exceeds a given detection threshold in its channel.

The results further validate the recurring idea in speech perception literature (Studdert-Kennedy, 1976) that there is a systematic difference in the processing of vowels and consonants. The validation stems from the fact that the difference between vowels and consonants appears consistently across various experimental paradigms mentioned in the introduction. Notably, the results for the stop, liquid, and vowel classes of speech sounds vary in a parallel manner to those in categorical perception. As the perception of stops, liquids, and vowels varies from most to least categorically perceived, so also the degree of interaction between the channels of analysis decreases correspondingly from stops through liquids to vowels. A result of a similar nature has been reported by Jamieson and Cheesman (1986) who found in a selective adaptation experiment that specific cues to voiced and voiceless consonants are processed by distinct auditory mechanisms located at different levels of the auditory system.

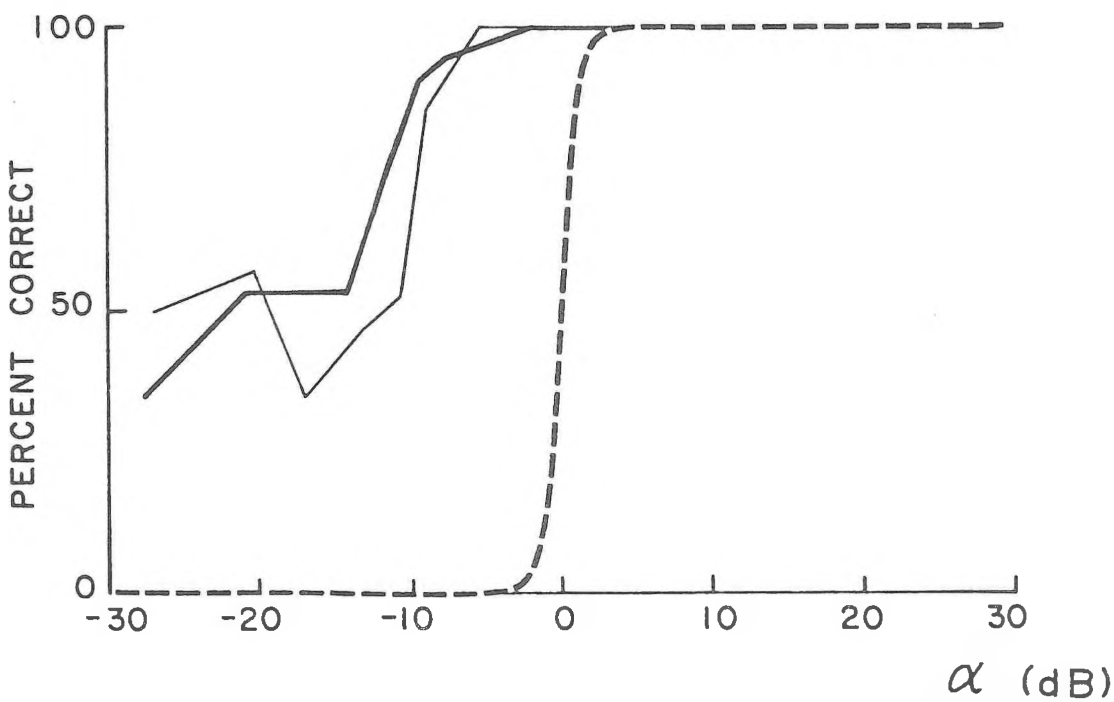


Figure 3. Identification by two subjects of the /bæ/ (heavy solid line) and /dæ/ (light solid line) stimuli in the presence of the /æ/ vowel masker, plotted as a function of the signal to masker amplitude ratio α expressed in decibels. As a reference, the averaged /bæ/-/dæ/ identification curve from Figure 1 is also shown (dashed line).

A note of caution should be made about the generality of the hypotheses of the dependence of the channels of auditory analysis for consonants and their independence for vowels. A simpler hypothesis can be based on peripheral properties of the auditory system, with acoustic similarity or dissimilarity of the mixed stimuli taken into account, irrespective of whether these stimuli are speech sounds or not. Obviously, if only the peripheral mechanisms were involved in the interaction between the mixed stimuli, it would be expected that simultaneous masking between two stimuli will become stronger as their spectral components become closer in frequency. To test whether this happens with our stimuli, a small-scale masking experiment was conducted. A masking stimulus was created by replicating one of the pitch periods of the steady-state vowel /æ/ to match the pitch contour of the /b/ formant transition. In addition, the amplitude of the /æ/ masker was scaled so that the intensities of the corresponding /æ/ and /b/ pitch periods were identical. This /æ/ vowel onset, when concatenated to the steady-state vowel from which it was gated out, produced a natural sounding vowel /æ/. This masker was combined with either the /b/ or /d/ formant transitions using 21 different relative amplitude weightings. Thus two sets of 21 stimuli were created, each ranging from pure /æ/ to either pure /bæ/ or /dæ/. These stimuli were presented in random order in a fully crossed design using an identical testing procedure to that described above. Two subjects, fully informed as to the nature of the experiment, were tested. Their task was to identify each stimulus as either "bæ" or "dæ". The percentages of correct identifications as a function of the amplitude weighting factor α are presented in Figure 3. For the purpose of comparison, the averaged /bæ/-/dæ/ identification curve from Figure 1 is also shown. The results indicate that the masking effect of the /æ/ component on the /b/ and /d/ transitions increased gradually with its amplitude: the correct identification curves are less steep than the identification curve from the /bæ/-/dæ/ experiment. Furthermore, for both the /b/ and /d/, nearly 100% correct identification was observed, even for stimuli which were about 10 dB weaker than the vowel masker. This indicates that the mutual masking between the /b/ and /d/ is more effective than that of the /æ/ masker, in spite of the fact that the frequency separation of loci for the F2 formant transitions for /b/ and /d/ is wider than for the vowel masker and either /b/ of /d/. This is not what would be expected if spectral masking were the only factor involved.

The authors are aware of the limited scope of this study. Before these hypotheses could be proposed for a wider class of speech stimuli, a fuller range of speech stimulus pairs must be tested. In the present study, only two phonetically similar consonant pairs and one phonetically dissimilar vowel pair have been tested. For greater confirmation of the hypotheses, two further tests must be carried out. One should include consonant pairs that vary in degrees of dissimilarity such as /ba/-/va/, /ba/-/wa/, or even /ta/-/sa/. Alignment of such stimuli can pose greater technical problems due to possible differences in temporal structure of these stimuli. The other test should include pairs of highly similar vowels such as /i/-/ɩ/. If the dissimilar consonant pairs still fused and the similar vowel pairs did not, only then would the results constitute strong evidence for the hypotheses suggested above.

The vowel combination /i/-/æ/ was the only one tested. Examination of other vowel pairs may prove to be productive, since varied perceptual effects can be expected, depending on the formant values of the two mixed vowels. Chiba and Kajiyama (1958, p.218) list these possible results: both vowels are clearly distinguished, only one of the two vowels is heard, either one or both of the vowels change quality, the two vowels are perceived as one vowel of different quality, or one of the vowels is heard in the background of the other.

By repeated playback of a boundary stimulus from a stop consonant continuum it is possible to hear either one or the other component of the mixed stimulus, as the listener so chooses, but never both components simultaneously. The receptor appears to be in a bistable state, where a change in the subjects' bias can shift the response from one category to another. This auditory effect can be likened to the "Necker cube" phenomenon in visual perception. Either of two forms of this reversible visual pattern, where all edges, including the "hidden" ones, are equally prominent, can be perceived, but never both simultaneously. The binaural counterpart of this phenomenon has been noted by Ades (1974) who observed that simultaneous presentation of /bæ/ into one ear and /dæ/ into the other results in a single fused percept heard as either one or the other stimulus. The "Necker cube" phenomenon further supports the hypothesis of inhibitory interaction between channels of phonetic analysis, in which the channels interact in such a way that only the most strongly excited one is perceived. Saunders (1980, p.95) noted that this phenomenon has three characteristic features of the cusp catastrophe: bimodality, sudden jumps, and hysteresis. Such sensory behavior was modelled by Zeeman (1976) using Duffing's equation.

The monaural superposition method was also tested with a three-way combination of stimuli (Stevenson, 1979). A mixed /ba/-/da/-/ga/ stimulus was constructed by component weighting α/b/ + β/d/ + γ/g/ concatenated with an /a/, where $\alpha+\beta+\gamma=1$. The scaling factors were varied in steps of 0.05. The stimuli were presented as in the previous experiments, except that three response options "ba", "da", and "ga" were provided. The results for the two subjects tested are shown in Figure 4, where the three symbols represent the modal values of ten judgements at each testing condition. The division of the response space into three distinct regions is apparent. The boundaries intersect in a trivalent point. By and large, the stimuli were perceived as clear exemplars of either /ba/, /da/ or /ga/.

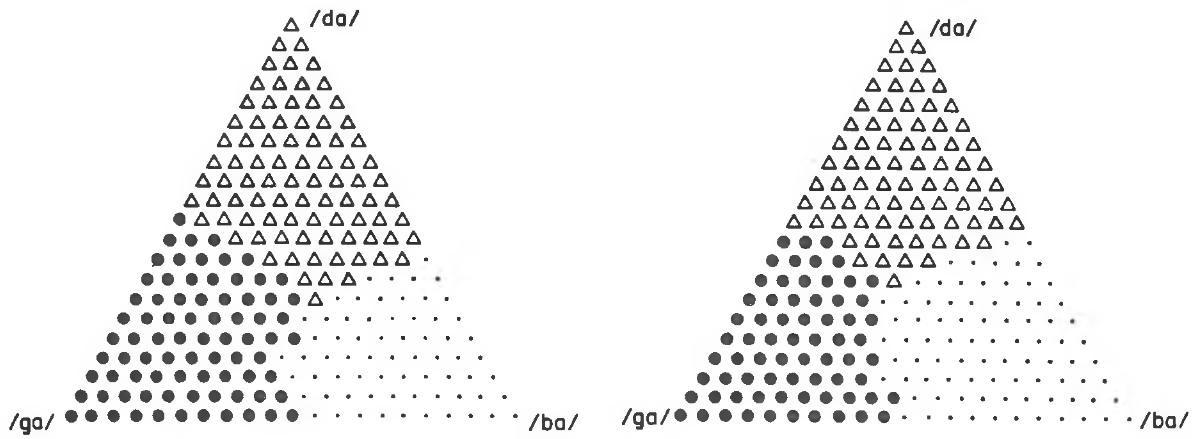


Figure 4. Results of the /ba/-/da/-/ga/ identification test for subjects JH (left) and DS (right).

APPENDIX: NOTES ON THE MONAURAL SUPPERPOSITION TECHNIQUE

The usefulness of the technique lies in the fact that real speech can be used to construct the stimuli, thereby preserving all necessary cues in the stimuli. Furthermore, identification of the endpoint stimuli on the continuum is unequivocal for all subjects. The monaural superposition technique can be equally well applied to synthetic stimuli with the advantage of simpler alignment of the mixed components, since they can be generated with an identical pitch contour and temporal pattern. Stimuli for multiple-category identification tasks can be generated by mixing more than two components. The monaural superposition stimuli are characterized by the amplitude weighting factor of the mixed components, representing a simple unidimensional continuum which permits an easy specification of the independent stimulus variable by a single physical parameter.

Although a variety of speech tokens can be used, a perfect monaural fusion can be obtained only with stimuli of similar temporal structure which differ only in spectral cues. The method works best with CV or CVC stimuli where the vowel is the same for both constituent stimuli and the mixed consonants, either in the initial or final position, are members of the same manner class, but differ in place of articulation. Within-category discrimination of stimuli for which monaural fusion takes place is increasingly difficult as the stimuli approach the category prototype. This property of the monaural superposition technique may make it unsuitable for discrimination experiments testing categorical perception. The reader is referred to Repp (1981) and Stevenson *et al.* (1985) for additional information on the applicability, advantages, and limitations of the monaural superposition technique

The monaural superposition technique merits comparison with the traditional technique of interpolated parameters used heretofore in testing categorical perception of speech stimuli. Judged on the basis of an identification experiment, the mixed stimulus continuum and the traditional continuum of interpolated parameters may appear as perceptually equivalent. Comparing these two continua using synthetic stimuli /da/-/ga/, Repp (1981) not only found that they produce virtually identical identification curves, but also that the increment in reaction time for boundary stimuli was of equal magnitude, suggesting that the boundary stimuli produced by the two techniques are equally ambiguous. The distinction between the two methods could emerge only if the subjects are allowed open or multiple-category responses such as confidence ratings of the responses, evaluations of the difficulty of the task, or naturalness judgements of the stimuli. The equivalence reported by Repp is not likely to be observed, for instance, on the /b/-/g/ continuum, provided the subjects are also permitted the "d" response. This continuum has been tested indirectly in the three-category identification experiment /ba/-/da/-/ga/ by the stimuli with the missing /da/ component, i.e., $\beta=0$. Corresponding results can be found on the bottom row of the response triangle in Figure 4. For both subjects tested, a complete absence of "d" responses was evident, even in the raw data. This precludes the hypothesis that a psychoacoustic fusion, i.e., an averaging of F2 transitions of /ba/ and /ga/, may produce a /da/-like percept (Cutting, 1976). On the other hand, in a corresponding interpolated parameters experiment, the F2 transition of the synthesized stimuli must necessarily pass through the /d/ locus of this formant, and thus be likely to yield some "d" responses. Radically different results between the two methods can be expected in the vowel perception experiments. While in the mixed /i/-/æ/ experiment both vowels are heard simultaneously, in a parallel interpolated parameters experiment the responses should cover the full range of the front vowels /i/-/ɪ/-/e/-/ɛ/-/æ/, heard one vowel at a time.

REFERENCES

Ades, A.E. (1974). "Bilateral component in speech perception," J. Acoust. Soc. Am. **56**, 610-616.

Ainsworth, W.A. (1977). "Mechanisms of selective feature adaptation," Percept. Psychophys. **21**, 365-370.

Chiba, T. and Kajiyama, M. (1958). *The Vowel: Its Nature and Structure* (Phonetic Society of Japan, Tokyo).

Cutting, J.E. (1975). "Aspects of phonological fusion," J. Exp. Psychol. Hum. Percept. Perform. **104**, 105-120.

Cutting, J.E. (1976). "Auditory and linguistic processes in speech perception: Interference from six fusions in dichotic listening," Psych. Rev. **83**, 114-140.

Eimas, P.D., and Corbit, J.D. (1973). "Selective adaptation of linguistic feature detectors," Cogn. Psychol. **4**, 99-109.

Elman, J.L. (1979). "Perceptual origins of the phoneme boundary effect and selective adaptation to speech: A signal detection theory analysis," J. Acoust. Soc. Am. **65**, 190-207.

Halwes, T.G. (1969). *Effects of Dichotic Fusion on the Perception of Speech,* Ph.D. thesis, University of Minnesota, Minneapolis, Minnesota.

Jamieson, D.G. and Cheesman, M.F. (1986). "Locus of selective adaptation in speech perception," J. Exp. Psychol. Hum. Percept. Perform. **12**, 286-294.

Miller, J.L. (1975). "Properties of feature detectors for speech: Evidence from the effects of selective adaptation on dichotic listening," Percept. Psychophys. **18**, 389-397.

Miller, J.L. (1977). "Properties for feature detectors for VOT: The voiceless channel of analysis," J. Acoust. Soc. Am. **62**, 641-648.

Repp, B.H. (1976). "Identification of dichotic fusions," J. Acoust. Soc. Am. **60**, 456-469.

Repp, B.H. (1980). "Stimulus dominance in fused dichotic syllables: Trouble for the category goodness hypothesis," J. Acoust. Soc. Am. **67**, 288-305.

Repp, B.H. (1981). "Perceptual equivalence of two kinds of ambiguous speech stimuli," Bull. Psychonom. Soc. **18**, 12-14.

Rozsypal, A.J. (1976). "Digital gating of speech signals," Lang. Speech **19**, 57-74.

Rozsypal, A.J., Stevenson, D.C., and Hogan, J.T. (1985). "Dispersion in models of categorical perception," J. Math. Psychol. **29**, 271-288.

Saunders, P.T. (1980). *An Introduction to Catastrophe Theory* (Cambridge University Press, Cambridge).

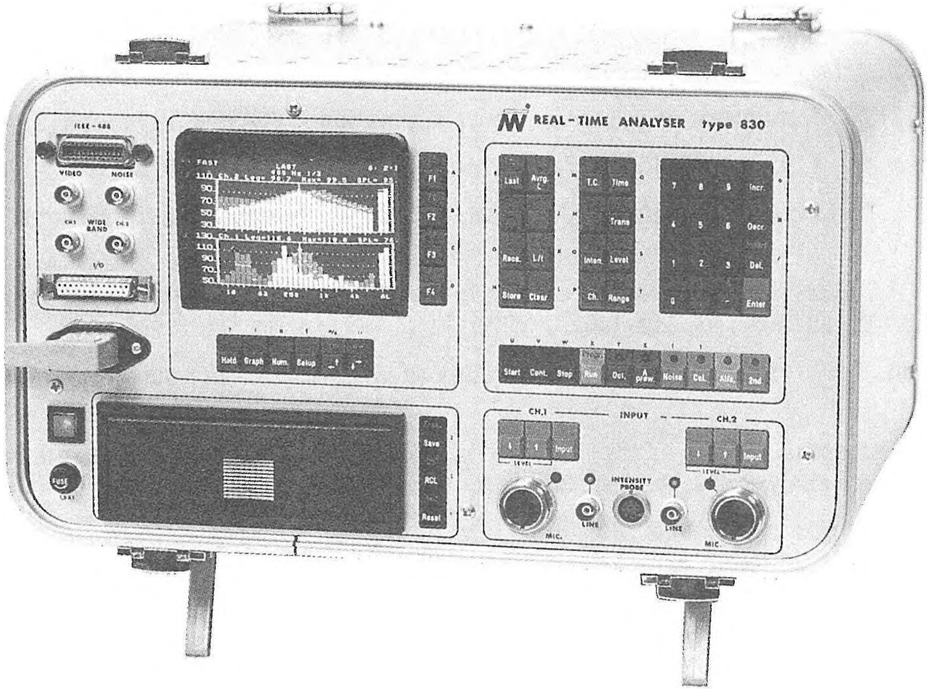
Stevenson, D.C., and Stephens, R.C. (1978). "A programming system for psychoacoustic experimentation," Proc. of the 11th DECUS (Canada) Symposium, Ottawa, 15-17.

Stevenson, D.C. (1979). "Categorical Perception and Selective Adaptation Phenomenna in Speech," Ph.D. thesis, University of Alberta, Edmonton, Alberta.

Stevenson, D.C., Hogan, J.T., and Rozsypal, A.J. (1985). "Generation of speech continua through monaural fusion," Behav. Res. Math. Instr. Comput. **17**, 102-106.

Studdert-Kennedy, M. (1976). "Speech perception," in *Contemporary Issues in Experimental Phonetics*, edited by N.J. Lass (Academic Press, New York), pp. 243-293.

Zeeman, E.C. (1976). "Duffing's equation in brain modelling," Bull. Inst. Math. and Appl. **12**, 207-214.



The Norwegian Electronics Type 830 Real-Time Analyzer is owned by many of the largest and highly regarded Acoustical Labs in the free-world

WHY NOT THE BEST?

The reasons are clear:

- ☑ Digital filters
- ☑ $^1/_1$ or $^1/_3$ octave analysis
- ☑ 80 dB dynamic range
- ☑ Internal storage
- ☑ Self contained
- ☑ Sound intensity
- ☑ Reverberation time
- ☑ IEEE-488 and RS-232 interfaces
- ☑ Internal Basic
- ☑ Upgradable
- ☑ 0,8 Hz to 20 kHz and higher
- ☑ True dual channel capability
- ☑ Sophisticated trigger facilities
- ☑ Small, rugged and portable
- ☑ Color video output
- ☑ Printer output
- ☑ 12 months warranty
- ☑ Support and service
- ☑ Future options
- ☑ Sound intensity mapping software

Contact us for more information

SCANTEK INC.

1559 Rockville Pike, Rockville, MD 20852. (301) 468-3502
Subsidary of Norwegian Electronics a.s. Representatives throughout North America.

State-of-the-art instrumentation