CANADIAN JOURNAL OF PSYCHOLOGY, 1985, 39(1), 1–15

Opponent Processes in Vowel Perception*

William F. Thompson, Lola L. Cuddy, and Barrie J. Frost
Queen's University

ABSTRACT Three experiments examined the possibility that the perception of vowels involves an opponent organization. Five vowel sounds were chosen on the basis of the vowel circles mapped by Yilmaz (1967, 1968). Whispered vowels were recorded and consisted of two pairs of complementary vowels (/u/(hoot) and /æ/ (hat), and /o/ (hoe) and /e/ (hay)) plus the most neutral English vowel (nearest to the centre of the vowel circle) /ɔ/(the). Experiment 1 revealed that mixtures of complementary vowels were most often confused with the neutral vowel /ɔ/. Experiment 2 revealed that complementary vowel mixtures and the neutral vowel /ɔ/ were rated most similar of all pair types excluding those in which the same vowel was present in both members of the pair. The final experiment examined the perception of pure and mixed vowels after adaptation to pure vowels. Analogous to adaptation in colour perception, adaptation to a vowel resulted in a perceptual shift of the neutral vowel and vowel mixtures towards the complementary vowel.

RÉSUMÉ Dans trois expériences, les auteurs étudient la possibilité que la perception des voyelles implique une organisation antagoniste. Cinq voyelles sonores sont choisies à partir des cercles de voyelles établis par Yilmaz (1967, 1968). Des voyelles chuchotées sont enregistrées. L'enregistrement comprend deux paires de voyelles complémentaires (u/(hoot) & /æ/ (hat) et /o/ (hoe) & /e/ (hay)), ainsi que la voyelle anglaise la plus neutre (la plus proche du centre du cercle de voyelle) /ɔ/ (the). L'expérience 1 montre que les mélanges de voyelles complémentaires sont le plus souvent confondus avec la voyelle neutre /ɔ/. L'expérience 2 révèle que les mélanges de voyelles complémentaires et la voyelle neutre /ɔ/ sont évalués comme les plus similaires de tous les types de paires, à l'exclusion de ceux dans lesquelles la même voyelle est présente dans les deux membres de la paire. Dans l'expérience finale, les auteurs étudient la perception des voyelles pures et des voyelles mixtes après une adaptation aux voyelles pures. Comme dans le cas de la perception des couleurs, l'adaptation à une voyelle produit un glissement perceptif de la voyelle neutre et des mélanges de voyelles vers la voyelle complémentaire.

One of the great insights into the physiology of vision was E.G. Hering's notion of opponent colours (Hering, 1878, 1920). From his observation that certain pairs of colours are never perceived as hue qualities at the same point in time and space, Hering called such colours "opponent" and inferred that a parallel opponency must also occur at the physiological level. The importance of opponent mechanisms was indicated by the discovery that individual nerve fibres often exhibit spontaneous discharge activity and respond to stimulation by either an

*This paper is based upon a thesis submitted by the first author in partial fulfillment of the M.A. degree at Queen's University in Kingston. The research was supported by the Natural Sciences and Engineering Research Council through a pre-doctoral fellowship awarded to the first author and an operating grant awarded to the second author. We thank Alan Marr for creating the software necessary for this research. Thanks are also due to C.L. Searle and to B.W. Tansley, who provided many helpful comments during the writing of this paper. Address reprint requests to L.L. Cuddy, Department of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6.

increase or a decrease in firing (e.g., Hartline, 1949; Hartline & Ratliff, 1957). Furthermore, Svaetichin, Negishi, and Fatehchand (1965) recognize Hering's opponent processes as "compatible with the standard neurophysiological principle of excitation-inhibition, which is operative for instance in a flexor-extensor control system" and consider the opponent colour system as a natural consequence of the functional design of the nervous system (Svaetichin et al., 1965, p. 179).

In light of the success of the concept of opponent mechanisms in explaining colour perception, it is not surprising to find this notion applied to the study of speech perception. Yilmaz (1967, 1968) invoked the notions of opponent mechanisms and sensory channelling in a theory of speech perception in which he claimed that there are three primary vowels from which all vowels can be constructed, and that there are complementary vowels just as we have complementary colours. Yilmaz derived a "vowel circle" by plotting the sine and cosine fourier components of the spectral distributions of vowel sounds. As with the colour circle, the vowel circle was argued to predict perceptual effects for additive mixing and adaptation. Importantly, the vowel map was designed to model what Yilmaz assumed to be physiological mechanisms underlying vowel perception: data reduction by the channelling of vowel information, and the operation of opponent mechanisms. Thus, while Hering had modelled phenomenological observations and inferred physiological processes, Yilmaz modelled what he considered to be widely applicable physiological processes and predicted phenomenological observations. The experiments reported here examine some of the predictions made by Yilmaz.

The use of gross spectral shape in the representation of speech has been advocated by Yilmaz as an alternative to formant-based representations (e.g., Peterson & Barney, 1952). However, the method of deriving the vowel map is not a critical issue for the present discussion. Searle (1982a, b) has pointed out the topographical similarity among the Yilmaz vowel circle, formant-based vowel maps (e.g., Slawson, 1982) and vowel maps recovered from principle-components analysis (Pols, 1977). What is of crucial importance, as Searle subsequently emphasizes, is the implication of the Yilmaz approach for modelling a general system of perceptual processing. Accordingly, the system separates relevant from irrelevant information and processes the relevant information through a minimum number of perceptually important channels or dimensions balanced around a single referent. At this level of organization, the principles involved in the perceptual processing of colour, such as complementarity, contrast, and transformation, have a general application (cf. Richards, 1979). In Searle's own model of auditory processing, the system of data reduction requires a level of neural representation subsequent to and operating upon the initial filter-bank analysis of basilar membrane activity.

It is also important to note that Yilmaz's approach is not contradictory to existing approaches in speech analysis and perception. Cutting (1976), for example, showed that a mixture composed of the formant structures of /ba/ and /ga/ yielded the percept /da/, which is not present in the spectral mixture, but is the result of data reduction in the form of perceptual averaging of the second formants. A mixture of vowels may be subject to the same averaging, such that a vowel of "midcolour" between the two inputs is perceived (Cutting, 1976, p. 135). While the same sort of averaging is predictable from Yilmaz's model, a key concept not discussed by Cutting is that of the neutral percept which, like the perception of white light, is not only the balance point of complementaries, but also a reference point about which the perceptual space is organized. For the perception of vowels or vowel-like sounds, the neutral in both the Yilmaz representation and phonetic maps (e.g., Gray, 1939) is the schwa or /ə/.

Because the concept of the neutral vowel is so critical to Yilmaz's approach to vowel perception, the experiments that follow test two basic predictions concerning its nature and operation. These predictions run as follows: (1) The vowel /ə/ (the neutral vowel) can be produced by a mixture of vowels whose position on the vowel circle is approximately opposite. The pairs tested in the present experiment were the vowels /e/ and /o/, and the vowels /u/ and /æ/. (2) The position of the neutral vowel is not invariant, but may be shifted by adapting to another vowel. The shift will be in a direction on the vowel circle opposite to the adapted vowel.

EXPERIMENT 1: MIXING COMPLEMENTARY VOWELS

The first experiment was conducted to determine whether mixtures of complementary vowels resemble the neutral vowel /ə/. Pilot work indicated that mixtures of voiced vowels produced clear auditory streaming (Bregman, 1978), and were most certainly heard as the two component vowels mixed together. Whispered vowel mixtures, however, were predicted to be unified since they would not contain the audible frequency beats contained in voiced vowel mixtures. A variety of mixtures of two complementary vowel pairs were chosen for the purpose of estimating the proportions of the mixtures that most closely resembled the neutral /ə/. Because the vowel circle is mapped slightly differently for each speaker, it was not assumed that a 50-50 mixture of complementary vowels would most resemble the neutral vowel. However, if Yilmaz's vowel circle model is a good description of vowel perception, then all mixtures of complementary vowels should tend toward the perception of the neutral vowel. Furthermore, a particular proportion of mixture components, although not necessarily one of the proportions used in the present experiment, should create a sound that is identical to the neutral vowel.

Method

Subjects: Eight men and eight women from Queen's University served as subjects in the experiment. The subjects volunteered their services and were naive concerning the purposes of the experiment. All participants had normal hearing and ranged in age between 15 and 35.

Apparatus and Stimuli: Signals were recorded and mixed via a DEC PDP 11/23 host computer and a DMX-1000 Digital Signal Processor (Digital Music Systems, Boston). Whispered vowels were recorded digitally on the host by a 1023AD analogue/digital converter from ADAC Corporation. The sampling rate of the recording was 19,305 samples per second. The sound was

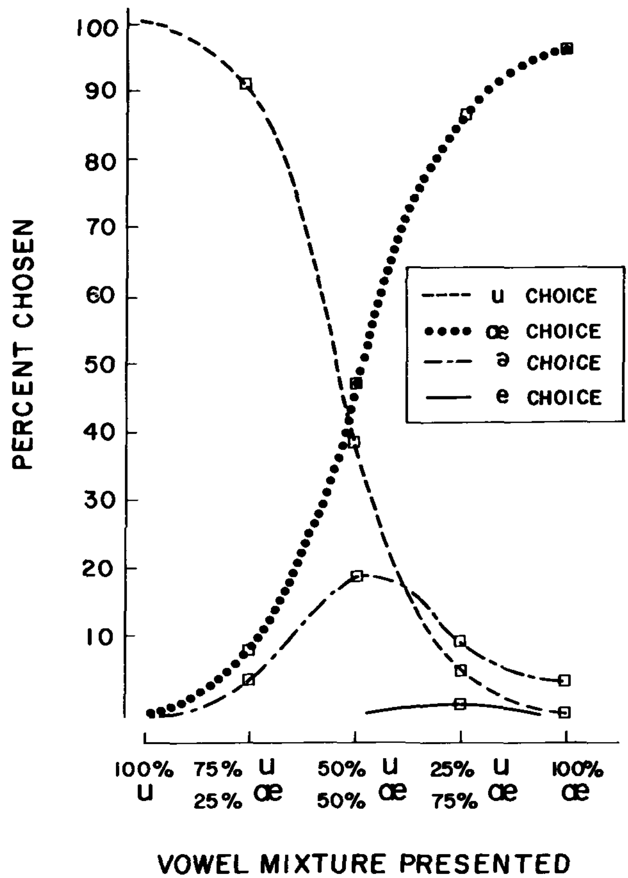

Figure 1. Listener's perception of /u/, /æ/, and three mixtures of /u/ and /æ/. (Percent /o/ choice, .1%, was too small to be included in the figure.)

filtered with a low pass filter at 6000 Hz with 24 dB per octave rolloff and a high pass filter at 20 Hz with 24 dB per octave rolloff. For each vowel, 12,800 samples were collected (approximately 660 msec duration) including the onset and decay of the vowel envelope. The samples of each vowel were scaled so that the overall RMS power of the vowels was equivalent. The front of the vowels was shaped by a sharp slope to remove a slight click at the onset caused by the threshold-triggering technique used in the digital recording. Recorded vowels were digitally mixed in a variety of proportions on the host computer. The experimenter adjusted the vowel onset times so that perceptual simultaneity occurred. Pure and mixed vowels were played back by the DMX-1000 signal processor and were recorded in a random order on Sony FeCr cassette tape. The tape was played on a Sony TC-K61 cassette deck. The sound level was adjusted to a comfortable listening level for each subject (usually about 70 dB SPL).

Procedure: Subjects were seated in a double-wall soundproof booth and listened to the tape through Sennheiser headphones. They were given two presentations of the five possible vowel sounds, followed by one block of practice trials and five blocks of test trials. Each block of trials consisted of a randomly ordered presentation of the following twelve vowel sounds: (1) three mixtures of the complementary vowels /u/ (hoot) and /æ/ (hat): (a) 25% /u/, 75% /æ/ (b) 50%

/u/, 50% /æ/ (c) 75% /u/, 25% /æ/;(2) three mixtures of the complementary vowels /o/ (hoe) and /e/ (hay): (a) 25% /o/, 75% /e/ (b) 50% /o/, 50% /e/ (c) 75% /o/, 25% /e/; (3) two examples of the pure (approximately) neutral vowel /ə/ (the); and (4) one example of each of the pure vowels, /u/, /e/, /o/, and /æ/.

After each presentation subjects were to choose from a list of five vowel categories (/æ/, /u/, /o/, /e/, or /ə/) the category that best corresponded to what they heard. Choices were made on response sheets provided by the experimenter. Subjects were asked to use their first impression when making a choice and were told that there were no right or wrong answers.

Results

Figure 1 displays responses to the various mixtures of the vowels /u/ and /æ/. When neither of the vowel components of the presentation was chosen, subjects chose the neutral vowel /ə/ 95% of the time. This frequency is significantly different from an equal response frequency of the three noncomponent vowels, $\chi^2 = 92.80, p < .001$. /ə/ responses were highest to the 50-50 mixture of /u/ and /æ/ where it was chosen approximately 18% of the time. The /ə/ response function is significantly different from a flat response function, $\chi^2 = 48.97$, $p < .001$.

Figure 2 displays responses to the various mixtures of the vowels /e/ and /o/. When neither of the vowel components of the presentation was chosen, subjects chose the neutral vowel /ə/ 60% of the time. This frequency is not significantly different from an equal response frequency of the three noncomponent vowels, $\chi^2 = 5.20$, $p > .05$. The percentage of /ə/ responses were highest to the 75% /o/, 25% /e/ mixture where it was chosen approximately 4.3% of the time. The /ə/ response function is significantly different from a flat response function, $\chi^2 = 20.44$, $p < .001$.

Discussion

Examination of the figures reveals that vowels presented in isolation were recognized with near 100% accuracy. This finding is interesting in view of previous work suggesting that the perception of isolated vowels is unreliable (e.g., Fairbanks & Grubb, 1961; Strange, Verbrugge, & Shankweiler, 1974). More importantly, the figures suggest that mixtures of complementary vowels not only resemble the components of the mixture, but also resemble the neutral vowel /ə/. Thus, the data yield modest support for Yilmaz's contention that the neutral vowel can be created by a mixture of opponent vowels.

Only three mixtures of each complementary pair were presented, and therefore it is possible that the ideal mixtures for optimal perception of the neutral vowel were not sampled. Consider especially the mixtures of /o/ and /e/, which were expected to yield a similar response function to that of /u/ and /æ/ mixtures. For the 50-50 /u-æ/ mixture, subjects chose both mixture components with approximately equal frequency. This mixture also had the highest neutral vowel response. None of the three /o-e/ mixtures resulted in the components /o/ and /e/ being chosen with near equal frequency, and none of these mixtures were heard as /ə/ nearly to the extent that the 50-50 /u-æ/ mixture was heard as /ə/. Thus, it is likely that an /o-e/ mixture yielding a near equal response frequency for its

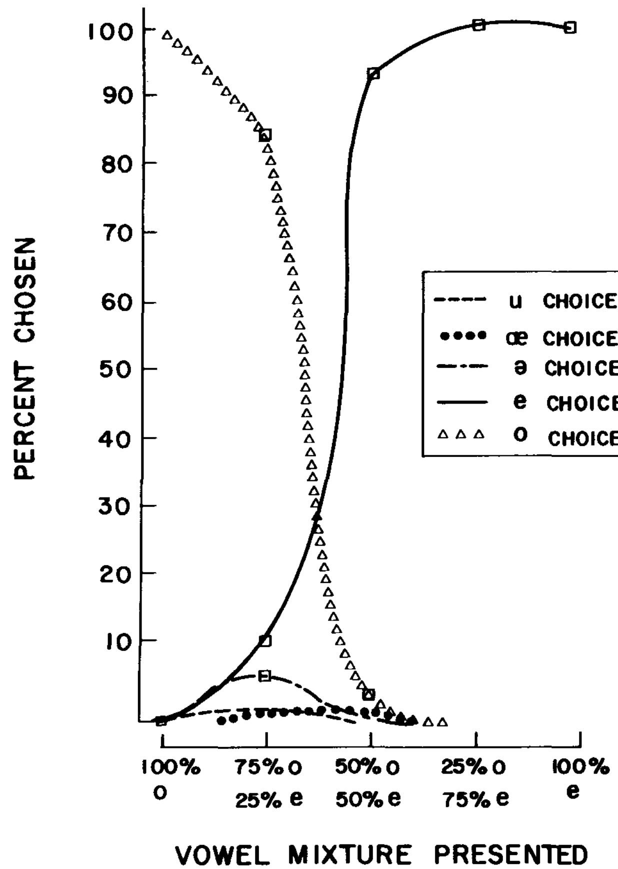

Figure 2. Listener's perception of /o/, /e/, and three mixtures of /o/ and /e/.

component vowels would also yield a higher percentage of neutral vowel percepts. Even so, it does not appear from the data that any mixture could produce the neutral vowel in the powerful way that mixtures of complementary colours produce the colour white. Cues that facilitate auditory streaming are clearly effective in isolating component vowels out of a mixture (Bregman, 1978; Bregman & Pinker, 1977; Dannenbring & Bregman, 1978; Steiger & Bregman, 1981).

Another problem is that the experimental paradigm did not allow a test of the following possibility: Although streaming may favour the choice of a component vowel to represent the percept of a mixture, the mixture may still yield a neutral vowel quality not present in any individual component. The second experiment was designed to test this possibility. In addition, the second experiment was intended to ensure that the /ə / choice was not merely a default choice representing the response "none of the above."

EXPERIMENT 2

The second experiment allowed subjects to rate the similarity between pairs of vowel sounds presented. If /ə/ had merely been a default category as suggested above, then it should not be given a high similarity rating to mixtures of complementary vowels. Further, the ratings should also reveal whether there is a *weak neutral quality present in complementary vowel mixtures but not in their* component vowels. Only one mixture of each complementary pair was used. Each mixture represented the median value, found by interpolation, of each of the /ə/ response functions of the first experiment. It was expected that the neutral vowel paired with mixtures of complementary vowels would be given higher similarity ratings than the neutral vowel paired with other pure vowels. It was also expected that complementary vowel mixtures would sound more like the neutral vowel than like any other noncomponent vowel of the mixture. In light of the results of the first experiment, it was expected that mixtures paired with a component of the mixture would be given very high similarity ratings.

Method

Subjects: Eight men and eight women from Queen's University served as subjects in the experiment. Subjects volunteered their services and were naive concerning the purposes of the experiment. Participants all had normal hearing, and were between the ages of 18 and 31.

Apparatus and Stimuli: Presentations were made from the analogue-digital recordings of the five pure vowels, and the vowel mixtures: (a) 50% /u/, 50% /æ/; and (b) 68% /o/, 32% /e/. Two blocks of 45 trials were recorded onto Sony FeCr cassette tape. Each trial consisted of a pair of vowel sounds separated by a 0.5 second pause.

Procedure: Subjects were seated in a sound proof booth and heard the tape through Sennheiser headphones. First they were presented with two exposures to the five pure vowels in order to familiarize them with the vowels used in the experiment. Subjects were then given 12 randomly selected practice presentations, followed by 2 blocks of 45 vowel-pair presentations. For each block the 45 vowel-pairs were given in a random order. The pair types are as follows: (1) 20 presentations of pure vowels paired with mixtures (i.e., 5 pure vowels $\times$ 2 mixtures $\times$ 2 orders); (2) 20 presentations of pure vowels paired with other pure vowels (using both orders for each pair); (3) five presentations of pure vowels paired with themselves.

After each stimulus pair was presented to the subject, she or he was required to rate the similarity between the two vowel sounds on a scale of 1 (least similar) to 7 (most similar). Subjects were instructed to disregard the clarity of the sounds and to judge only how easily the sounds could be thought of as being from the same vowel category.

Results

Of major interest in testing the validity of the Yilmaz vowel projection was a comparison of mean similarity ratings for three basic pair types. These three pair types are as follows: (1) complementary vowel mixtures (/u-æ/ or /o-e/) with the neutral vowel /ə/; (2) complementary vowel mixtures with the other pure vowels that were noncomponents of the mixture; and (3) the neutral vowel /ə/ with other pure vowels (/u/, /æ/, /e/, /o/).

As illustrated in Figure 3, the first pair type was rated significantly higher than either the second or the third pair types, $F(1, 15) = 29.02, p < .001$. The second

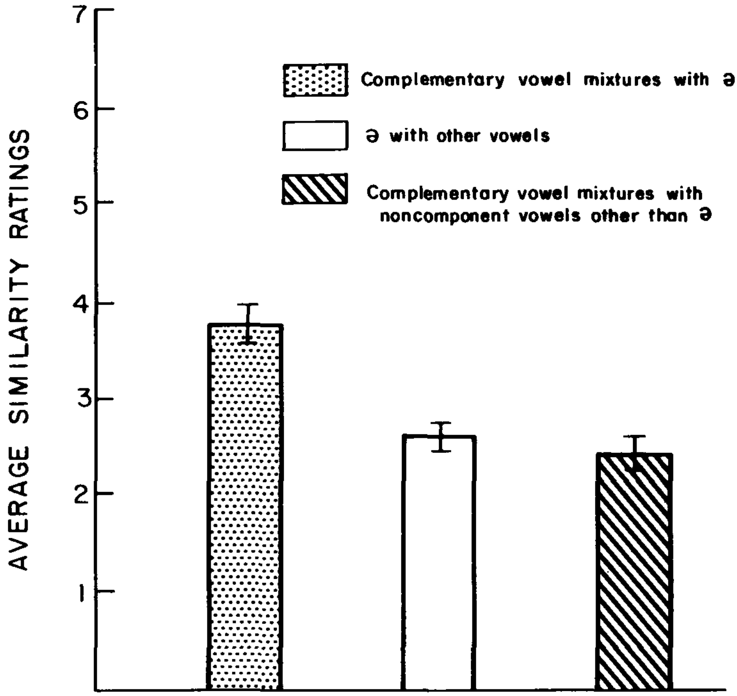

Figure 3. Average similarity ratings under three pair type conditions. Vertical lines indicate standard errors for each condition.

and third pair types were not rated significantly differently from each other, $F(1, 15) = 0.51, p = .51$.

Pure vowels paired with themselves were always rated as maximally similar and were not considered in the analysis. The highest ratings considered for analysis were for complementary vowel mixtures paired with one of the vowels that made up the mixture. Ratings of this pair type were significantly higher than ratings for all other pair types, $F(1, 15) = 216.50, p < .001$. The lowest ratings were for pure vowels paired with other pure vowels. Ratings of this pair type were significantly lower than ratings of all other pair types, $F(1, 15) = 172.72, p < .001$.

The 68:32 /o-e/ mixture paired with the neutral vowel was given significantly higher similarity ratings than was the 50:50 /u-æ/ mixture paired with the neutral vowel, $F(1, 15) = 20.48, p < .01$. While this finding was not expected, it is not contrary to the vowel circle model.

Discussion

The results suggest that /ə/ was not a default category in Experiment 1 but actually does resemble mixtures of complementary vowels. Aside from mixtures paired with their component vowels, mixtures paired with /ə/ were given the highest similarity ratings. That the second and third basic pair types were not rated significantly differently from each other is also predictable by the theory: Since /ə/ and complementary vowel mixtures closely resemble each other, vowels paired with either should produce similar ratings. That the 68:32 mixture of /o/ and /e/ was rated as more similar to the neutral than the 50:50 mixture of /u/ and /æ/ gives us reason to suspect that the best mixture of /o/ and /e/ was not, in fact, used in Experiment 1, but was closely approximated in Experiment 2.

Similarity ratings of complementary vowel mixtures paired with their component vowels were very high. This finding is important: It is not analogous to the perception of complementary colour mixtures in vision. In colour perception, component colours are not seen in complementary colour mixtures. Possibly, the vowel mixtures did not mix perceptually because the vowel sounds were not in a speech context. It is also very possible that in any context of the stimuli, an opponent process in speech perception could never be clearly demonstrable by the methods used for colour perception. Complementary vowel mixtures may not be equivalent to the neutral vowel, but may nevertheless contain a neutral vowel quality that is virtually absent in the component vowels of the mixture.

EXPERIMENT 3: ADAPTING TO VOWELS

It is apparent from the first two experiments that mixtures of complementary vowels sound mostly like their component vowels, but also seem to have a neutral vowel quality. If an opponent process is involved, as Yilmaz has suggested, an adaptation effect should be demonstrable also. The present experiment can be distinguished from previous speech adaptation work in a few important ways. First, the selective effects of adaptation have usually been thought to suggest a fatiguing of selectively tuned feature detectors, and adaptation periods have been necessarily long (e.g., Diehl, 1975; Eimas & Corbit, 1973; Tartter & Eimas, 1975). An opponent process, however, is not to be confused with neuronal fatigue, and an adaptation period of less than 10 seconds should yield an effect (see Bailey, 1974, cited in Moore, 1982). Second, while most speech adaptation work concerns the issue of phonetic versus nonphonetic feature detectors (e.g., Morse, Kass, & Turkienicz, 1976; Roberts & Summerfield, 1981), the present experiment addresses general auditory principles which are not restricted to either phonetic or nonphonetic material. Finally, the concept of a neutral, which is critical to Yilmaz's model, is not merely a boundary between two categorical percepts (e.g., Eimas & Corbit, 1973), or a middle category on a particular continuum (e.g., Cooper, 1974; Morse et al., 1976), or the perceptual average of acoustic features (Cutting, 1976); it is all of the above, and also acts as a referent about which the entire perceptual space is organized.

The third experiment was designed to test the perception of pure and mixed vowel sounds after a brief adaptation to a pure vowel. Because of the length of each trial and the nature of the rating task devised, it was decided that only one complementary vowel pair would be used along with the neutral vowel in the present experiment. It was expected that adaptation to a pure vowel would result in a perceptual shift of the test vowel sound towards the complement of the adapted vowel. Such a shift was expected to be most detectable for test vowel mixtures that resembled the neutral (/u-æ/ and /u-ə-æ/) and for the pure neutral test vowel /ə/.

Method

Subjects: Eight men and eight women from Queen's University volunteered to be subjects (some had participated in Experiment 2). All subjects were naive concerning the purposes of the experiment, had normal hearing, and were between the ages of 18 and 35.

Apparatus and Stimuli: From the analogue-digital recordings of the pure vowels (1) /ə/, (2) /u/, (3) /æ/, and equal mixtures of (4) /u/ and /æ/, (5) /ə/ and /u/, (6) /ə/ and /æ/, and (7) /ə/, /u/, and /æ/, 3 blocks of 28 trials were recorded onto Ampex reel to reel tape. Each trial consisted of a pure vowel adaptation period followed by a single test sound. The tape was spliced such that for each recorded trial, the time interval between the adaptation period and the test sound was equal to the time interval between the repeated vowels of the adaptation period. The finished product was then recorded onto Sony FeCr cassette tape.

Procedure: Subjects were seated in a soundproof booth and heard the tape through headphones. After two exposures to the three pure vowels that were to be considered for rating, subjects were presented one practice block followed by two experimental blocks. For each of the 28 trials in a block, subjects were given 18 seconds to rate the test sound. The test sound was one of the three pure vowels or one of the four mixed vowel sounds (listed above) and was presented under the following four conditions: (1) immediately following adaptation to /æ/; (2) immediately following adaptation to /u/; (3) immediately following adaptation to /ə/; and (4) without adaptation.

Adaptation consisted of 12 rapid repetitions of the vowel. The time interval between repeated vowels was approximately 0.5 seconds. Subjects were instructed to consider only the last vowel sound for those presentations consisting of more than one vowel sound. After each presentation subjects were asked to give three presence ratings, corresponding to the relative presence of the three possible vowel sounds /ə/, /u/, and /æ/. This paradigm allowed subjects the option of reporting the perception of mixtures as more than one vowel. Ratings for each vowel sound were scaled from 1 (not heard) to 7 (predominant).

Results

Figure 4 shows average presence ratings for /æ/ and /u/ when the test sound was the mixture /u-æ/ (panels A and B) or the pure vowel /ə/ (panels C and D). Ratings given to the test complementary vowel mixture /u-æ/ were significantly different when preceded by /u/ adaptation than when preceded by /æ/ adaptation. For this test sound, presence ratings of /æ/ were higher following /u/ adaptation than following /æ/ adaptation, $F(1, 15) = 27.81$, $p < .001$. Also, presence ratings of /u/ were higher following /æ/ adaptation than following /u/ adaptation, $F(1, 15) = 8.45$, $p < .05$. As expected, no such differences were found for the presence ratings of /ə/, $F(1, 15) = 0.74$, $p = .59$.

Presence ratings given to the neutral test vowel were also found to differ under the varying conditions of adaptation. This difference occurred with the /æ/ ratings, which were higher following /u/ adaptation than following /æ/ adaptation, $F(1, 15) = 8.96$, $p < .01$. Presence ratings of /u/ tended to be very categorical (either 1 or 7) and were not significantly higher under conditions of /æ/ adaptation, $F(1, 15) = 1.0$, $p = .34$.

Ratings given to the mixture containing all three vowels /u-ə-æ/ were also expected to differ when preceded by /u/ adaptation than when preceded by /æ/

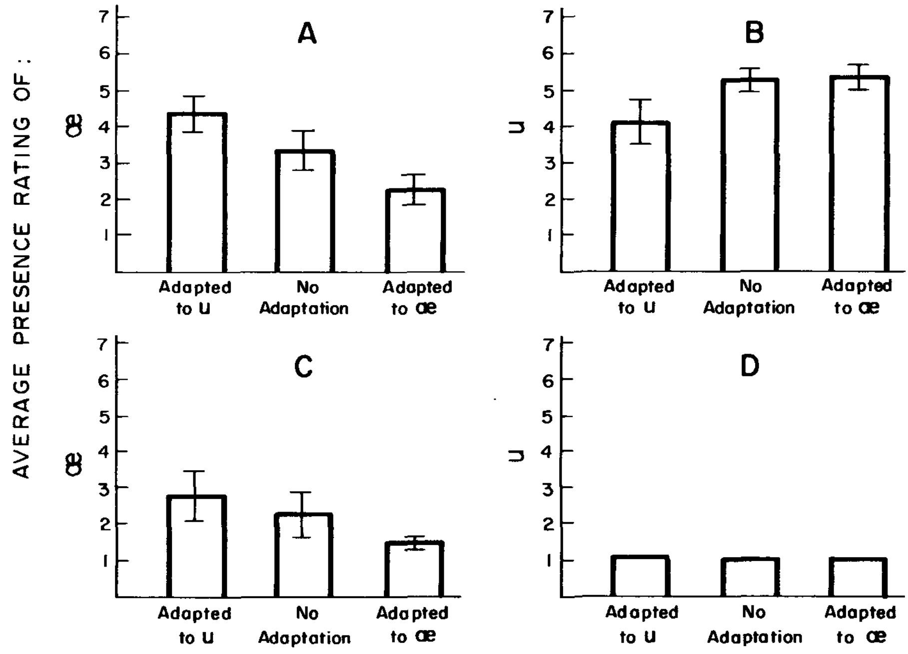

Figure 4. (a) Presence ratings of /u-æ/ for the test vowel mixture /u-æ/ under three adaptation conditions. (b) Presence ratings of /u/ for the test vowel mixture /u-æ/ under three adaptation conditions. (c) Presence ratings of /æ/ for the test vowel /ə/ under three adaptation conditions. (d) Presence ratings of /u/ for the test vowel /ə/ under three adaptation conditions. Vertical lines indicate standard errors for each condition. All three standard errors in condition (d) were zero.

adaptation. Though not displayed in the figure, presence ratings of /æ/ were again higher following /u/ adaptation than following /æ/ adaptation, $F(1, 15) = 11.7$, $p < .01$; and presence ratings of /u/ were higher following /æ/ adaptation than following /u/ adaptation, $F(1, 15) = 19.84, p < .001$. Presence ratings of /ə/ were unexpectedly found to differ also, depending on the adaptation, and were higher under conditions of /u/ adaptation, $F(1, 15) = 6.51, p < .05$.

Perceptual shifts due to adaptation were only towards the complement of the adapted vowel. For instance, adaptation to /u/ resulted in no shift in the perception of the test sound /æ/ toward /ə/.

Supporting the findings of Experiments 1 and 2, /ə/ was rated as present in 75.8% of the presentations of the test vowel mixture /u-æ/. /ə/ was rated as present in the complementary vowel mixture /u-æ/ significantly more often than /æ/ was rated as present in the vowel mixture /ə-u/, $\chi^2 = 17.20, p < .001$. /ə/ was also rated as present in the mixture /u-æ/ significantly more often than /u/ was rated as present in the vowel mixture /æ-ə/, $\chi^2 = 118.60, p < .001$.

Discussion

Even with the small amount of adaptation used in the experiment, adaptive shifts were detectable for the perception of test vowel mixtures and detectable in one direction for the perception of the pure neutral test vowel (towards /æ/ after /u/ adaptation). Adaptation to /æ/ was not able to produce a detectable shift in the perception of the pure test vowel /ə/ towards /u/, because subjects tended to rate /u/ very categorically (either absent or maximally present). In terms of a vowel circle, the /u/ recorded may lie at a peripheral position on the circle and far away from the recorded neutral vowel /ə/. That adaptation to /u/ was able to shift the perception of /ə/ detectably towards /æ/ suggests that /ə/ and /æ/ lie in closer positions on the vowel circle. This notion is supported by results obtained in Experiment 2 whereby the vowel pair /æ-ə/ was given significantly higher similarity ratings than the vowel pair /u-ə/, $F(1, 15) = 8.0, p < .05$. This finding is also consistent with the vowel mapping of Shepard (1974), which is based on confusions among vowels.

Although /u/ was rated categorically when the test sound was a pure vowel, subjects were more flexible when the test sound was a vowel mixture. This finding is not surprising since, for all adaptation conditions, subjects tended to give the component vowels of test vowel mixtures noncategorical ratings (not 1 or 7). Subjects were consequently not hesitant to use the entire range of noncategorical rating scores, thus enabling the detection of slight perceptual shifts.

GENERAL DISCUSSION

In the first experiment, presentations of certain vowel mixtures were heard reasonably often as the neutral vowel /ə/, but seldom as other noncomponent vowels of the mixture. On the other hand, pure vowels were recognized with near 100% accuracy. These findings were taken to support the notion that mixtures of opponent vowels have a neutral (or /ə/-like) quality that is not a quality of the pure vowels that compose the mixtures. However, since mixtures were heard primarily as one of their component vowels, it is apparent that testing Yilmaz's claim that there is a functional analogy between colour perception and vowel perception is not a simple task.

The second experiment used similarity ratings as an alternative test of the neutral quality of vowel mixtures. Omitting vowels paired with themselves, and mixtures paired with one of their components, the highest average similarity rating was found to be for mixtures of complementary vowels paired with the neutral vowel /ə/. These findings indicate that /ə/ was not used merely as a default category in Experiment 1 and that mixing complementary vowels does seem to create a third neutral vowel quality that is not perceived in pure vowels. The analogy to colour perception is apparent, although the equivalence between mixtures of complementaries and the neutral is much more striking in colour perception. Finally, the results of Experiment 3 are consistent with the involvement of an opponent process in vowel perception, in that performance shifts were obtained only in a direction away from the adapting vowel toward its complement.

Earlier, we introduced two proposals made by Yilmaz: All vowels are processed through the activity of a limited number of vowel channels; and the perception of vowels is organized by opponent mechanisms around a neutral reference point. If we accept the assumption that the neutral vowel /ə/ is analogous to achromatic light and that only two pure vowels are sufficient to create a percept of the neutral vowel, we may infer that the vowel continuum is sampled by a minimum of three vowel channels (Richards, 1979). More focal to our discussion, the data are consistent with the concept of opponent mechanisms. In this respect, mixtures of opponent vowels yielded a (partial) cancellation of the two vowel qualities, and hence were heard with an added neutral vowel quality. That the analogy between the neutral vowel and achromatic light is but a partial one is not contrary to Yilmaz's proposals. Yilmaz carefully acknowledged the differences as well as the similarities between visual and auditory processing. Colour information is synthesized at a very early stage of visual processing; the vowel map and its synthetic properties reflect a level of processing beyond an initial stage of multi-channelled analysis. It is not surprising, therefore, to find only partial evidence of opponent mechanisms for vowel perception with the present paradigm. However, if the assumption that /ə/ is analogous to achromatic light is rejected, Experiments 1 and 2 at least demonstrate the possibility of additive vowel mixing. Additive vowel mixing is predicted by Yilmaz's first proposal, though a strong case for this claim would require a demonstration that different vowels can be produced by different vowel mixtures.

Because Yilmaz intended to develop a system of general principles encompassing a variety of approaches in speech perception, his own approach does not create predictions that are necessarily incompatible with other approaches. However, the data do rule out certain alternative explanations. For instance, it might be proposed that the whispered neutral /ə/ and the complementary mixtures were judged to be similar merely because they were both uninterpretable noise. However, the neutral /ə/ was accurately identified as a pure vowel (Exp. 1) and given consistent presence ratings of /ə/ (Exp. 3), while mixtures were reliably streamed into two components plus a neutral quality. Both stimuli were therefore interpretable. Further, a fast-fourier-transform analysis of the stimulus sounds revealed visibly discernable formant structures for all test sounds in line with what would be expected from previous analyses of vowel sounds. In particular, spectral analyses of mixtures did not resemble each other nor did they resemble the spectral analysis of the neutral vowel. Thus, we attribute the judged similarity between the neutral vowel and complementary vowel mixtures to a process of data reduction, rather than a spectral similarity.

An alternative account of Experiment 3 is that adaptation induced a criterion bias such that responses leaned towards any nonadapted vowel in the test stimulus. While such an account can explain the ratings of complementary vowel mixtures which followed adaptation to one of the two component vowels, it cannot explain why adaptation to /u/ caused the perception of /ə/ to shift towards /æ/, but not /æ/ towards /ə/, or why when the test sound /u-ə-æ/ followed adaptation to /u/ or to /æ/, the perceptual shift was not equally divided between the two nonadapted components of the mixture, but favoured the complementary vowel.

The present findings suggest that opponent processes, which have great explanatory value in colour perception, may also play a role in the perception of vowels. Importantly, this work has also revealed some difficulties in drawing an analogy between colour perception and speech perception. While complementary vowel mixtures did contain a neutral vowel quality, they were not equivalent to the neutral vowel, and we are not certain that refinements in vowel mixing would necessarily achieve such an equivalence. Auditory perception is highly susceptible to analytic (streaming) effects, and it may be impossible to obtain experimental effects that are untouched by these processes. Nonetheless, exploring the effects of vowel mixing and vowel adaptation in the manner we have done may provide a useful tool for further investigation in this area, and it has forced us to consider that Hering's insight into opponent processes may apply to speech as well as vision.

REFERENCES

Bailey, P.J. (1974). Procedural variables in speech adaptation. *Speech Perception*, **2**, 29–34 (Queen's University of Belfast).

Bregman, A.S. (1978). Auditory streaming: Competition among alternative organizations. *Perception and Psychophysics*, **23**, 391–398.

Bregman, A.S., & Pinker, S. (1977). Auditory streaming and the building of timbre. *Canadian Journal of Psychology*, **32**, 19–31.

Cooper, W.E. (1974). Contingent feature analysis in speech perception. *Perception and Psychophysics*, **16**, 201–204.

Cutting, J.E. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review*, **83**, 114–140.

Dannenbring, G.L., & Bregman, A.S. (1978). Streaming vs. fusion of sinusoidal components of complex waves. *Perception and Psychophysics*, **24**, 369–376.

Diehl, R. (1975). The effect of selective adaptation on the identification of speech sounds. *Perception and Psychophysics*, **17**, 48–52.

Eimas, P.D., & Corbit, J.D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, **4**, 99–109.

Fairbanks, G., & Grubb, P.A. (1961). A psychophysical investigation of vowel formants. *Journal of Speech and Hearing Research*, **4**, 203–219.

Gray, L.H. (1939). *Foundations of language*. New York: Macmillan Company.

Hartline, H.K. (1949). Inhibition of activity of visual receptors by illuminating nearby retinal areas in the Limulus eye. *Federal Proceedings*, **8**, 69.

Hartline, H.K., & Ratliff, F. (1957). Inhibitory interactions of receptor units in the eye of Limulus. *Journal of General Physiology*, **40**, 357–376.

Hering, E. (1878). *Zur lehre vom lichtsinne*. Wien: Carl Gerald's Suhn.

Hering, E. (1920). *Grundzuge der lehre vom lichtsinn*. Berlin: Julius Springer.

Moore, B.C.J. (1982). *An introduction to the psychology of hearing* (2nd edition). New York: Academic Press.

Morse, P.A., Kass, J.E., & Turkienicz, R. (1976). Selective adaptation of vowels. *Perception and Psychophysics*, **19**, 137–143.

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, **24**, 175–184.

Pols, L.C.W. (1977). Spectral analysis and identification of Dutch vowels in monosyllabic words. Soesterberg, The Netherlands: Institute for Perception, TNO.

Richards, W. (1979). Quantifying sensory channels: Generalizing colorimetry to orientation and texture, touch, and tones. *Sensory Processes*, **3**, 207–229.

Roberts, M., & Summerfield, Q. (1981). Audio-visual adaptation in speech perception. *Perception and Psychophysics*, **30**, 309–314.
Searle, C.L. (1982a). Speech perception from an auditory and visual viewpoint. *Canadian Journal of Psychology*, **36**, 402–419.
Searle, C.L. (1982b). Some perceptual aspects of timbre. *Canadian University Music Review*, **3**, 80–101.
Shepard, R.N. (1974). Psychological representation of speech sounds. In E. David & P.B. Denes (Eds.), *Human communication: A unified view.* Maidenhead, England: McGraw-Hill.
Slawson, W.A. (1982). The musical control of sound color. *Canadian University Music Review*, **3**, 67–79.
Steiger, H., & Bregman, A.S. (1981). Capturing frequency components of glided tones: Frequency separation, orientation and alignment. *Perception and Psychophysics*, **30**, 425–435.
Strange, W., Verbrugge, R., & Shankweiler, D. (1974). *Consonant environment specifies vowel identity.* (Status Report on Speech Research, SR-37/38, 209–216). New Haven, CT: Haskins Laboratories.
Svaetichin, G., Negishi, K., & Fatehchand, R. (1965). Cellular mechanisms of a Young-Hering visual system. In A.V.S. de Reuck & J. Knight (Eds.), *Color vision: Physiology of experimental psychology.* Boston: Little, Brown and Company.
Tartter, V.C., & Eimas, P.D. (1975). The role of auditory and phonetic feature detectors in the perception of speech. *Perception and Psychophysics*, **18**, 293–298.
Yilmaz, H. (1967). A theory of speech perception. *Bulletin Of Mathematical Biophysics*, **29**, 793–823.
Yilmaz, H. (1968). A theory of speech perception: II. *Bulletin Of Mathematical Biophysics*, **30**, 455–479.

Received 15 February 1983
Revision received 16 February 1984
Accepted with revision 1 May 1984
Accepted 9 July 1984