CATEGORICAL PERCEPTION: THE GROUNDWORK OF COGNITION,
Stevan Harnad (Ed.), Cambridge University Press, 1987

8 Categorical partition: A fuzzy-logical model of categorization behavior

Dominic W. Massaro

The phenomenon of categorical perception (CP) is assessed by drawing upon a broad range of methodological, theoretical, and experimental issues. It is argued that the identification/discrimination task provides no support for the existence of CP. The categorical model usually provides an inadequate description of the results, and it has not been shown to provide a better description than alternative models. Even if the results in the task provided unequivocal support for the categorical model, alternative explanations are possible. It is necessary to distinguish between sensory and decision processes in the categorization task. Decision processes can transform continuous sensory information into results usually taken to reflect CP; and finding relatively categorical partitioning of a set of stimuli in no way implies that these stimuli were perceived categorically. Three new results are relevant to the issue of categorical versus continuous perception. First, subjects are asked to make repeated ratings of the degree to which a stimulus represents a given category. The distribution of the rating judgments to a given stimulus is more adequately described by a continuous than a categorical model of perception. Second, subjects asked to classify speech events independently varying along two dimensions produce identification results consistent with the assumption of continuous information along each of the two dimensions. A model based on categorical information along each dimension gives a very poor description of the identification judgments. Third, the reaction times of identification judgments illustrate that members within a speech category vary in ambiguity or the degree to which they represent the category. The best conclusion is to relabel [illegible] categorical partition [illegible] and to concentrate [illegible] world of speech.

[illegible]

accounts for the categorization of our stimulus world, while also accounting for our parallel ability to perform fine within-category discrimination or identification.

Categorical perception

In this chapter, I evaluate categorization processes in the domain of auditory-information processing, asking whether certain auditory continua are perceived categorically or in a continuous manner. Continuous perception refers to a relatively continuous relationship between changes in a stimulus and changes in the perceptual experience of that stimulus. In contrast, categorical perception refers to a mode of perception in which changes along a stimulus continuum are not perceived continuously, but in a discrete manner (Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). For example, listeners might be limited in their ability to discriminate differences between different speech sounds that belong to the same phoneme category. The sounds within a category can only be identified absolutely, and discrimination is possible for only those sounds that are identified as belonging to different categories.

Seminal study

In a seminal study, Liberman, Harris, Hoffman, and Griffith (1957) used synthetic speech to generate a series of 14 consonant–vowel syllables going from /be/ to /de/ to /ge/ (/e/ as in *gate*). The onset frequency of the second-formant transition of the initial consonant was changed in equal steps to produce the continuum. In the identification task, observers identified random presentations of the sounds as /b/, /d/, or /g/. The discrimination task used the *ABX* paradigm. Three stimuli were presented in the order *ABX*; *A* and *B* always differed and *X* was identical to either *A* or *B*. Observers were instructed to indicate whether *X* was the same as *A* or *B*. This judgment was supposedly based on auditory discrimination in that observers were instructed to use whatever auditory differences they could perceive.

The experiment by Liberman et al. (1957) was designed to test the hypothesis that listeners can discriminate stimuli only to the extent that they can recognize them as different phoneme categories. The hypothesis was quantified in order to predict discrimination performance from identification judgments. According to [illegible], stimuli can be discriminated only to the extent that they are identified as different. Figure 8.1 gives the results of a single subject for identification and discrimination reprinted from the article by Liberman et al. (1957). The authors concluded that the discrimination results were fairly well described by the predictions based on the identification judgments. This rough correspondence between identification and discrimination has provided the major source of support for the CP concept. In the most recent review of categorical-speech perception, the conclusion from this same study was that "the perception of these syllable-initial stops was invariably quite categorical" (Repp, 1984, p. 282). As Macmillan pointed out

---

The research reported in this paper was supported in part by grants from the National Institute of Mental Health and the National Institute of Neurological and Communicative Disorders and Stroke. I would like to thank Michael Cohen for help at all stages of the project and Neil Macmillan and Bruno Repp for comments on this chapter.

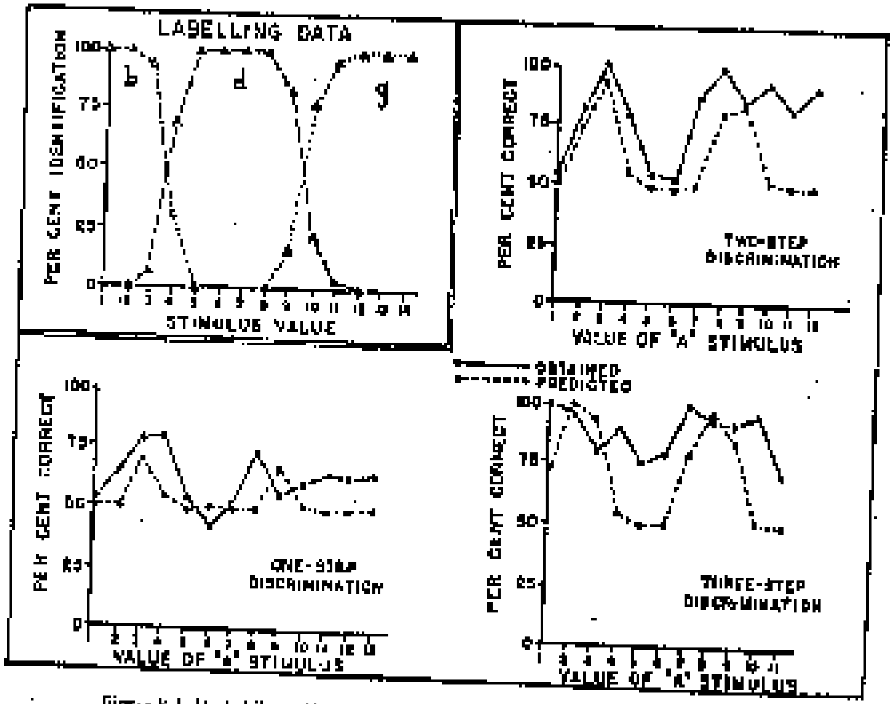

Figure 8.1. Probability of identification in the labelling task as a function of the 14 levels along the speech continuum (top left panel). The other three panels give the obtained and predicted probability of discrimination in the ABX task. The three panels correspond to different numbers of steps between the A and B stimulus. (After Liberman et al., 1957.)

(see Chapter 2, this volume), however, observed discrimination is usually better than that predicted by identification. For some reason, the discrepancy has never been a deterrent for advocates of CP, or a central result for the alternative view.

To provide a proper assessment of the CP model, it is necessary to determine how closely the predicted discrimination matches the observed and to compare this prediction with other possible predictions. The most direct goodness-of-fit measure is the root-mean-squared deviation (RMSD), which is the square root of the average squared difference between the predicted and observed points. The RMSD can take on values between zero and one if the predicted and observed values lie between zero and one. The rough measure of goodness of fit is simply how close to zero [illegible] RMSD values. In reality, the RMSD value [illegible] close to [illegible] values between zero and [illegible]. The results [illegible]

nation values are between 0.5 and 1 given the two-alternative task, a model predicting 75% correct discrimination across all conditions would have to give RMSD values of less than 0.25. With hindsight, we see that quantitative tests of the categorical model and just about any alternative model would have precluded the authors and others from finding support for the existence of CP.

Although performance in the identification/discrimination task already appears to be inadequately described by CP, the arguments that I support in this chapter disqualify the identification/discrimination task as diagnostic of CP. It is necessary to distinguish between the phenomenon of "categorical perception" and "categorical-perception results" in the identification/discrimination task. "Categorical perception" is defined in terms of whether the sensory system makes available continuous or categorical information, whereas "categorical-perception results" refer to the accuracy with which identification predicts discrimination. There are alternative explanations of CP results that do not depend on CP. I now turn to three possible explanations of this kind.

Auditory-memory limitations

The well-known contribution of memory limitations in the discrimination task can lead to CP results (Fujisaki & Kawashima, 1969; Massaro, 1975a,b; Papp, 1975; Pisoni & Lazarus, 1974). The failure to discriminate two sounds may simply reflect a limitation in short-term auditory memory for the sounds. Given the good evidence for the important role of auditory memory in the discrimination task (Papp, 1975; Pisoni, 1973; Pisoni & Lazarus, 1974), I will not discuss this evidence here. To the extent that short-term auditory memory is not available in the discrimination task, it is impossible for discrimination performance to exceed that predicted from identification.

I would emphasize that the present description of the role of auditory and verbal memory in the discrimination task differs from the dual-coding model first proposed by Fujisaki and Kawashima (1969). In the dual-coding model, verbal codes are primary and auditory memory is secondary in that it is only used if the stimuli to be discriminated are given the same verbal code. As nicely demonstrated by Macmillan (Chapter 2, this volume), the dual-coding model fails to predict discrimination performance. In my view (Massaro, 1976), however, subjects use auditory information in the discrimination task to the extent that it is present, and they may rely on verbal codes when the auditory information is insufficient. This interpretation of discrimination performance is not easy to test quantitatively, as the contribution of auditory memory cannot be measured in the traditional identification/discrimination task. Demonstrating the accuracy of the interpretation, however, is not [illegible] that the ability of identification to predict discrimination performance may be due to poor auditory memory and the use of verbal [illegible]

Qualitative dimensions

Results consistent with CP might also occur whenever a single acoustic continuum gives rise to both quantitative and qualitative perceptual dimensions. For example, Pastore et al. (1977) asked subjects to identify and discriminate loudness differences. When test tones of different amplitudes were presented in silence, the relationship between identification and discrimination was not consistent with the predictions of CP. With a reference tone present, however, the test tones *could* be perceived as small increases or decreases in amplitude relative to the reference tone. The reference tone caused a sharp boundary in the identification function and a corresponding peak in the discrimination function, as predicted by CP. Listeners were able to discriminate whether the test tone was louder or softer than the reference tone, but they were not able to discriminate the magnitude of the difference. Accordingly, sounds perceived as continuous in silence produced CP results in the context of a reference tone. Discriminating the direction of loudness change with respect to the reference tone overshadowed the continuous discrimination of absolute loudness. Hence, a qualitative dimension might be responsible for the apparent categorical perception of a quantitative dimension.

Context influences

Even without auditory-memory limitations and the contribution of qualitative dimensions in the identification/discrimination task, a correspondence between identification and discrimination is not sufficient evidence for CP. Contextual variables can always serve to encourage CP results, even though perception is continuous. The major goal of the experiments reported by Hary and Massaro (1982) was to demonstrate that a stimulus continuum can appear to be categorically perceived even though the sounds within each category along the continuum are noticeably different from one another. Achieving this goal would demonstrate that CP results in the identification/discrimination task do not necessarily mean that perception is categorical. If CP results can be found with a continuum of sounds shown to be perceived continuously, the identification/discrimination task must be rejected as a valid measure of CP.

By varying the amplitude rise time of tones, it is possible to generate a continuum of musical sounds that vary in their pluckish/bowish quality (Cutting & Rosner, 1974). The pluckish quality resembles a stringed instrument that is plucked, whereas the bowish quality resembles a stringed instrument that is bowed. Hary and Massaro (1982) studied the pluck/bow distinction but with the rise-time continuum extended in positive and negative directions. In the standard continuum, the rise times were varied in just the positive direction. In the bipolar continuum, the rise times varied in both negative and positive directions. Figure 8.2 illustrates the first 100 msec of the nine sounds along the bipolar continuum. The standard continuum consisted of the same five stimuli with positive rise times and four additional stimuli with positive rise times. All subjects were tested with both the bipolar and the standard continuum. Given that five of the sounds were identical in the two continua, it was possible to evaluate discrimination and identification of the same stimuli when tested within both types of continua. Based on the earlier results of Rosen and Howell (1981), Hary and Massaro (1982) predicted that the standard continuum would not produce categorical results. In addition, the presence of the positive and negative rise times along the bipolar continuum should lead to categorical results. Given continuous results along the standard continuum of rise times, the categorical results along the bipolar continuum would support the hypothesis that traditional CP results can be found along a continuum of sounds that is perceived continuously.

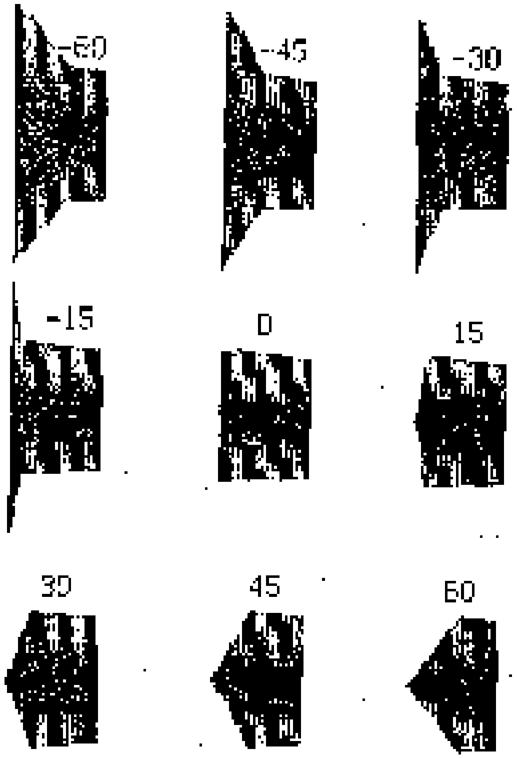

Figure 8.2. Illustration of the waveform of the first 100 msec of each of the nine sounds along the bipolar rise time continuum. The rise time is given at the top of each of the sounds. Negative numbers refer to negative rise times and positive numbers to positive rise times. The rise times of the nine sounds along the standard continuum ranged between 0 and 120 msec in steps of 15 msec.

Figure 8.3 gives the average identification, percentage correct discrimination,

Qualitative dimensions

Results consistent with CP might also occur whenever a single acoustic continuum gives rise to both quantitative and qualitative perceptual dimensions. For example, Pastore et al. (1977) asked subjects to identify and discriminate loudness differences. When test tones of different amplitudes were presented in silence, the relationship between identification and discrimination was not consistent with the predictions of CP. With a reference tone present, however, the test tones could be perceived as small increases or decreases in amplitude relative to the reference tone. The reference tone caused a sharp boundary in the identification function and a corresponding peak in the discrimination function, as predicted by CP. Listeners were able to discriminate whether the test tone was louder or softer than the reference tone, but they were not able to discriminate the magnitude of the difference. Accordingly, sounds perceived as continuous in silence produced CP results in the context of a reference tone. Discriminating the direction of loudness change with respect to the reference tone overshadowed the continuous discrimination of absolute loudness. Hence, a qualitative dimension might be responsible for the apparent categorical perception of a quantitative dimension.

Context influences

Even without auditory-memory limitations and the contribution of qualitative dimensions in the identification/discrimination task, a correspondence between identification and discrimination is not sufficient evidence for CP. Contextual variables can always serve to encourage CP results, even though perception is continuous. The major goal of the experiments reported by Hary and Massaro (1982) was to demonstrate that a stimulus continuum can appear to be categorically perceived even though the sounds within each category along the continuum are noticeably different from one another. Achieving this goal would demonstrate that CP results in the identification/discrimination task do not necessarily mean that perception is categorical. If CP results can be found with a continuum of sounds shown to be perceived continuously, the identification/discrimination task must be rejected as a valid measure of CP.

By varying the amplitude rise time of tones, it is possible to generate a continuum of musical sounds that vary in their pluckish/bowish quality (Cutting & Rosner, 1974). The pluckish quality resembles a stringed instrument that is plucked, whereas the bowish quality resembles a stringed instrument that is bowed. Hary and Massaro (1982) studied the pluck/bow distinction but with the rise-time continuum extended in positive and negative directions. In the standard continuum, the rise times were varied in just the positive direction. In the bipolar continuum, the rise times varied in both negative and positive directions. Figure 8.2 illustrates the first 100 msec of the nine sounds along the bipolar continuum. The standard continuum consisted of the same five stimuli with positive rise times and four additional stimuli with positive rise times. All subjects were tested with both the bipolar and the standard continuum. Given that five of the sounds were identical in the two continua, it was possible to evaluate discrimination and identification of the same stimuli when tested within both types of continua. Based on the earlier results of Rosen and Howell (1981), Hary and Massaro (1982) predicted that the standard continuum would not produce categorical results. In addition, the presence of the positive and negative rise times along the bipolar continuum should lead to categorical results. Given continuous results along the standard continuum of rise times, the categorical results along the bipolar continuum would support the hypothesis that traditional CP results can be found along a continuum of sounds that is perceived continuously.

Figure 8.3 gives the average identification, percentage correct discrimination,

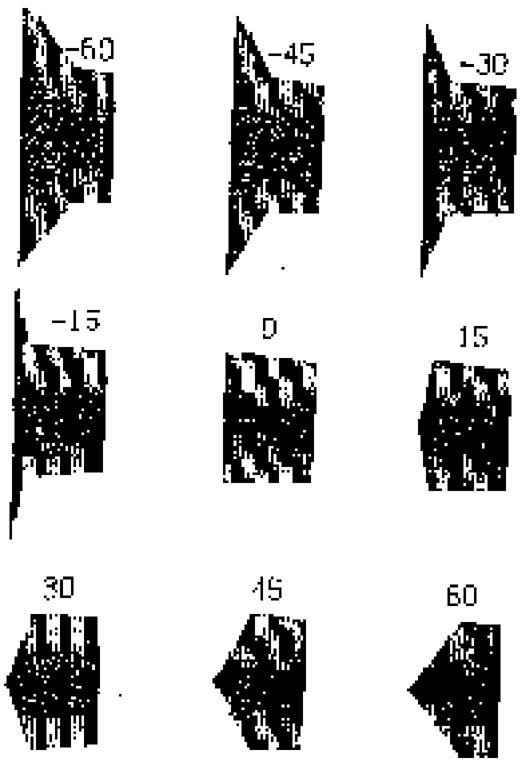

Figure 8.2. Illustration of the waveforms of the first 100 msec of each of the nine sounds along the bipolar rise-time continuum. The rise time is given at the top of each of the sounds. Negative numbers refer to negative rise times and positive numbers to positive rise times. The rise times of the nine sounds along the standard continuum ranged between 0 and 120 msec in steps of 15 msec.

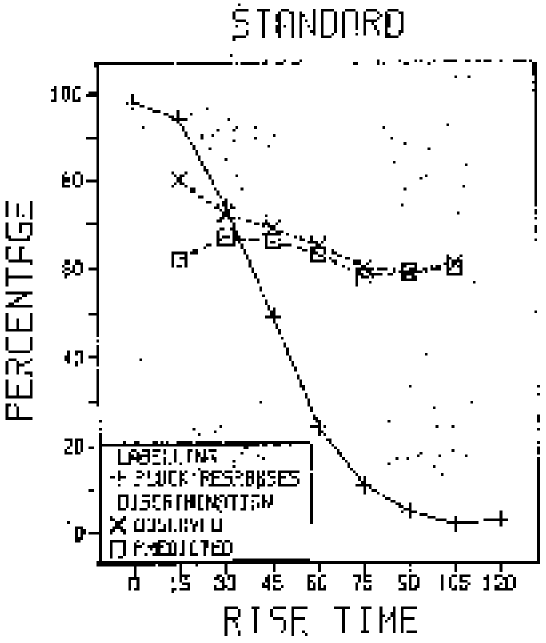

Figure 8.3. Percentage of "pluck" responses in the identification task as a function of the rise time of the test sound. Observed versus predicted percentage of correct discriminations in the *AX* discrimination task for the standard continuum. (After Hary & Massaro, 1982.)

and predicted discrimination functions for the standard condition. These functions are very similar to those obtained earlier by Rosen and Howell (1981). Both the identification and discrimination functions indicate that the subjects perceived the sounds along the continuum in a continuous manner. The identification function reveals a relatively gradual change in the percentage of "pluck" responses with increases in rise time. The discrimination function shows that subjects discriminated the 0- to 30-msec comparison most accurately, with accurate performance falling off continuously as rise time increased. These results from the standard continuum seem to be most consistent with continuous perception and the applicability of Weber's law to the rise-time continuum (Cutting, 1982; Rosen & Howell, 1981).

Figure 8.4 presents the average identification, discrimination, and predicted discrimination functions for the bipolar condition. The identification function shows a very sharp drop in the percentage of pluck responses between the 0- and 15-msec rise times. In addition, discrimination was best near the category boundary and relatively poor within the categories. Figure 8.4 shows that discrimination seems to be accurately predicted by identification performance. Normally these results would be taken as evidence of CP for these stimuli. However, the same subjects were able to discriminate the stimuli within a category more accurately when they were tested within the context of the standard continuum. It is clear that subjects could perceive within-category differences even though the bipolar results, considered alone, would traditionally lead to the conclusion of CP.

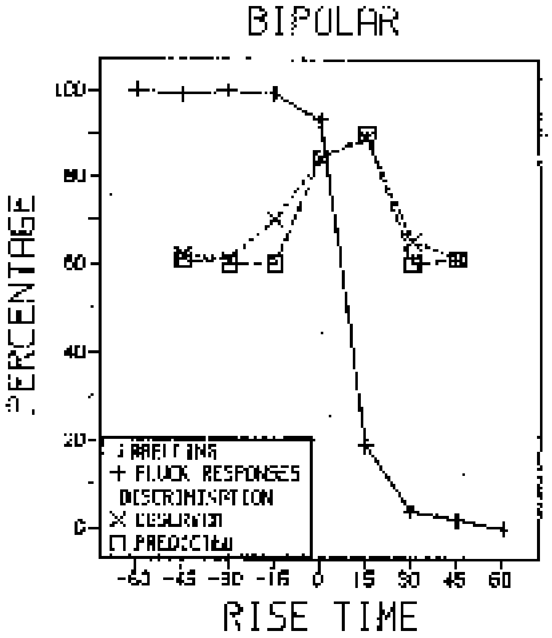

Figure 8.4. Percentage of "pluck" responses in the identification task as a function of the rise time of the test sound. Observed versus predicted percentage of correct discriminations in the *AX* discrimination task for the bipolar continuum. (After Hary & Massaro, 1982.)

Summary

Previous studies taken as evidence for CP have failed to demonstrate that the categorical model gave a good description of the results or a significantly better one than alternative models. In addition, categorical results in the identification/discrimination task can occur because of short-term auditory memory limitations, the presence of qualitative dimensions, and context influences. Thus, finding relatively good discrimination between categories and poor discrimination within a category or an agreement between identification and discrimination performance cannot be taken as evidence for CP. I now turn to a new framework and test of categorical versus continuous speech perception.

Sensory and decision processes in categorization

The issue of categorical perception might be clarified by using a model of performance derived from the field of psychophysics. It is now generally accepted that a

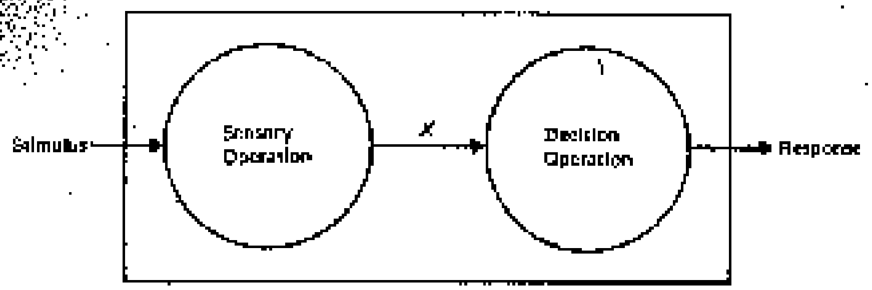

Figure 8.5. A flow diagram of two operations that occur between stimulus and response in an identification or categorization task. The sensory operation transmits some information, X, to the decision operation.

psychophysical task contains both sensory and decision stages of processing. Figure 8.5, taken from Massaro (1975a), illustrates the sequence of processing in a psychophysical task. Applying this model to the domain of CP, we can ask whether the output of the sensory stage before the decision stage is categorical. If it is, we call it CP. If it is only the output of the decision stage that is categorical, then we say that perception is continuous. Performance in the latter case might be described as "categorical partitioning" because the decision stage is usually required to make categorical responses. (Compare the pandemonium model in Chapter 6, by Remez and analog/digital filtering in Chapter 19, by Harnad, both in this volume.)

Continuous sensory information

A continuous model assumes that the sensory information is continuous, with the understanding that the meaning of a message is usually conveyed by discrete alternatives. The perceiver has continuous information about the degree to which a given alternative was presented, but must decide which pattern category was presented. Accordingly, the perceiver requires a decision algorithm to map continuous information into a discrete choice. It seems reasonable to assume that the listener establishes a criterion value of information and responds as a function of the observed value of information relative to the criterion value. This assumption is a direct application of the decision mechanisms given by the theory of signal detectability (Green & Swets, 1966).

Strong evidence for continuous perception of acoustic features comes from a training study carried out by Pisoni, Aslin, Perey, and Hennessy (1982). They taught native English speakers to categorize three rather than two categories across the voice-onset-time continuum. That subjects were able to learn the new category in just a short training session validates the assumption of continuous feature information. This experiment is similar in some respects to an experiment carried out a decade earlier by Barclay (1972) in which subjects were required to partition a

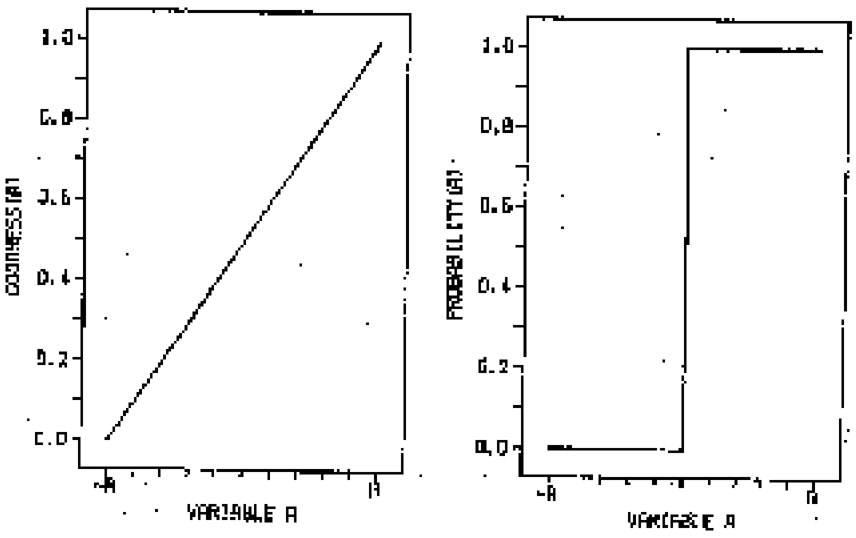

Figure 8.6. *Left panel:* Hypothetical category "goodness," Goodness(A), values as a function of the continuous changes in stimulus variable A. *Right panel:* The probability of an A response, probability(A), as a function of the continuous changes in stimulus variable A.

/b/-/d/-/g/ continuum into just two categories, /b/ and /g/. The categorization of a stimulus usually identified as /d/ was related to its similarity to the /b/ and /g/ stimuli. That is, subjects were more likely to identify a /d/ stimulus as /b/ to the extent it was similar to /b/. In both types of studies, subjects are able to modify their categorization of a stimulus continuum, and this categorization behavior must result from the differential partitioning of the continuous information along the speech dimension of interest.

Molfese (Chapter 14, this volume) reviews a series of experiments that recorded auditory-evoked responses (AER) to temporal cues in speech and nonspeech stimuli. Various components of the AER recorded from different sites proved to be sensitive to changes in the temporal-onset asynchrony between components of stop consonants and complex tones. The major finding relevant to our concerns is that the AER components consistently reflect stimulus differences within a category just about as well as, if not better than, stimulus differences between categories. This physiological evidence complements the behavioral evidence for continuous rather than categorical sensory information in both speech and nonspeech domains.

Discrete decision rule

Assume a stimulus continuum, A, that is perceived continuously. The value, $G(A)$, is defined as the goodness of A and is an index of the degree to which the information represents a given category. The left panel of Figure 8.6 presents a hypothetical function of the relative goodness of A, defined G(A), as a linear function of variable A. What is the most reasonable decision rule? An optimal decision rule would set the criterion value at 0.5 and classify the pattern as A for any value greater than this. Otherwise, the pattern is classified as non-A. In this case, the probability of an A response, P(A), would take the step function form shown in the right panel of Figure 8.6. That is, with a fixed criterion value and no variability, the decision operation changes the continuous linear function given by the operation into a step function. A step function occurs as long as the observer holds a fixed criterion value, regardless of the actual value held. It is worth noting that this function is also identical to the idealized form of CP in a speech-identification task (Studdert-Kennedy et al., 1970). Thus, a step function cannot be taken as evidence for CP because it can occur given continuous featural information.

Noise

A step function is rarely observed in identification, for what seem to be good reasons. If there is noise in the perceptual operation, a given level of variable A cannot be expected to produce the same G(A) on each presentation. If the noise is assumed to have equal variance across the stimulus continuum, then the noise near the criterion value will have a much larger consequence than the noise away from the criterion value. Figure 8.7 illustrates the outcome of pattern classification for two types of noise given a criterion value of 0.5. A safe assumption is that the noise is symmetrical at the mean, the mean is equal across the continuum, and the variance is equal across the continuum. A stimulus whose mean goodness, G(A), is at the criterion value will produce random classifications if the noise is symmetrical. With symmetrical noise, the value of G(A) will be above the criterion on half of the trials and below the criterion on the other half. As the mean perception moves away from the criterion value, the noise will have a diminishing effect on the classification response. Noise has a larger consequence at the middle of the range of G(A) values than at the extremes because variability goes in both directions in the middle and only inward at the extremes. If the noise is also normal, the resulting identification function will be ogival in form.

This decision algorithm allows categorical decisions to be made on the basis of continuous information. In addition, the model can predict classification functions with sharp boundaries, previously taken to represent CP. Some investigators (Siegel & Siegel, 1977a,b) have interpreted a sharp boundary along the identification function as evidence for CP. However, as I made clear in my formalization of a continuous perception model, a decision operation intervenes between the [illegible]

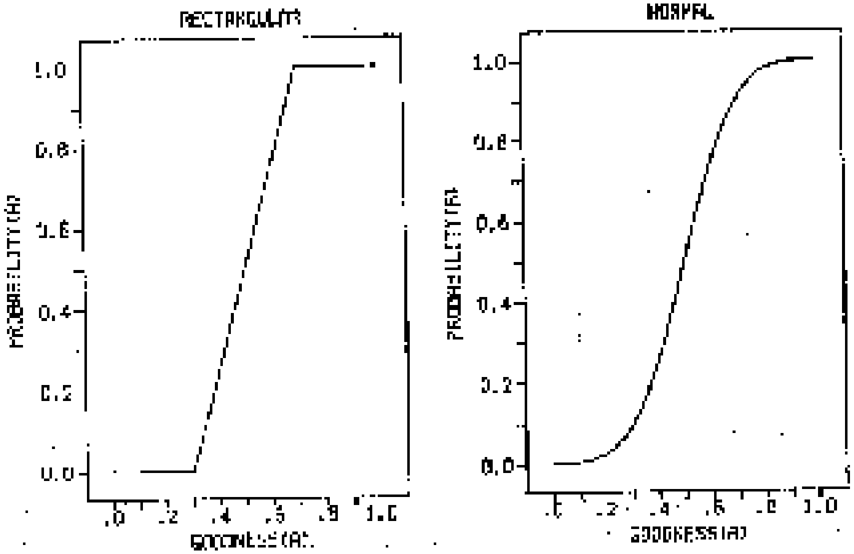

Figure 8.7. The predicted outcome of pattern classification given continuous perception of some variable, A, and two types of noise during the perceptual operation. The probability of an A identification is plotted as a function of the degree to which variable A is perceived as representing the category A, Goodness(A). The left panel assumes rectangular noise and the right panel, normal noise present during the perceptual operation.

New views of current explanations

This representation of the problem equates CP with threshold models and continuous perception with the theory of signal detectability. Because threshold models are currently out of favor, I might be accused of setting up a straw man rather than providing a fair representation of CP. However, one of the most attractive properties for supporters of CP has been that it stands in marked contrast to how perception usually works. In the domain of speech perception, some supporters have argued that CP was one characteristic that provided evidence for the qualitative difference between speech and nonspeech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

The two-stage model illustrated in Figure 8.5 also clarifies alternative explanations of CP. Explanations either in terms of auditory discontinuities (Pastore et al., 1977) or acoustic invariance (Stevens & Blumstein, 1978) are clearly statements about the output of the sensory stage. These views represent instances of threshold [illegible]

[illegible] in terms of speech-specific phonetic processing [illegible] For example, Repp and Liberman [illegible]

(Chapter 3, this volume) state that "the link between perception and production (in most general terms) enables the category prototypes to respond appropriately to articulatory or co-articulatory adjustments." However, an intelligent categorization at the decision stage, such as the one proposed by Repp and Liberman, is not sufficient to qualify speech as special because it occurs in so many other domains of categorization (Massaro, 1984; Oden, 1981).

Repp and Liberman interpret flexible boundaries in categorization as evidence for a phonetic mode rather than a pure auditory mode in speech perception. However, they fail to address the important issue of the relationship between flexible boundaries and categorical versus continuous information. I claim that flexible boundaries are more compatible with continuous than with categorical sensory information.

A new test

Given the distinction between sensory and decision processes, Massaro and Cohen (1983a) offered a new approach to the question of CP. The distinguishing feature of this approach is the use of continuous rather than discrete perceptual judgments (Watson, Rilling, & Bourbon, 1964). Although rating judgments have been used in a number of previous studies (Elman, 1979; Sawusch, 1977), they have not been analyzed so as to test between categorical and continuous models of speech perception. Relative to discrete judgments, continuous judgments may provide a more direct measure of the listener's perceptual experience. For example, McNabb (1974) found that a binary response proved insensitive to the manipulation of an independent variable whereas confidence ratings revealed significant effects of this variable. In the Massaro and Cohen (1983a) study, subjects were asked to rate the degree to which they felt that the speech stimulus represented one alternative or the other, rather than simply indicating which alternative was present. Categorical and continuous models of speech perception were formalized and evaluated against the distribution of repeated rating responses to each test stimulus along a speech continuum.

Categorical model

Consider the assumptions of the CP model illustrated in Figure 8.8. One is that the listener has only two perceptual classes or categories along a sound continuum of five levels. Referring to Figure 8.8, the output of the sensory stage can take on only the values *a* or *b*. At stimulus level 1, the likelihood of an *a* percept is very high whereas the likelihood of a *b* percept is very low. As the levels increase, the relative likelihood of the two percepts changes so that at level 5, *b* is the most likely percept. However, in all cases, the sound is heard as either *a* or *b*. If perception is truly categorical, any sound along the continuum can be heard only as *a* or *b* and nothing in between. What does a CP subject do when asked to make continuous rating judgments? He might note the foolishness of the request, but most probably would attempt to comply in a reasonable manner. The subject would choose a rating toward the *A* end of the response scale for the perception of *a* and toward the *B* end for the perception of *b*. If there is variability in memory and response, however, the subject would generate a distribution of rating responses for each of the two percepts. That is, subjects may not remember where they last rated the *a* category and they may also only approximate the intended rating because of response variability. Furthermore, given the demand characteristics of the task, subjects might actually generate additional variability in their ratings if their percepts were categorical and they were expected to make a range of rating responses. A safe assumption is that the rating responses to the percept *a* would be normally distributed, with a mean *Xa* and a variance *Sa* and in a similar manner for the *b* percept.

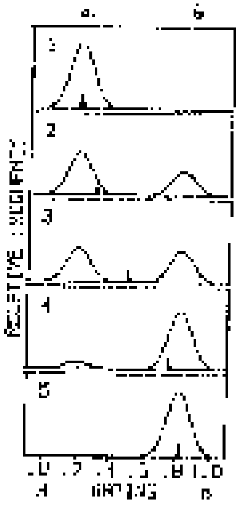

Figure 8.8 Illustration of the categorical-perception rating model. The sizes of the *a* and *b* distributions depict the likelihood of an *a* or *b* percept given each of five levels (1 through 5) along a hypothetical-stimulus continuum. The location and range of each distribution gives the rating responses that will be observed. The downward-pointing arrows give the mean ratings across the five levels of the stimulus continuum.

The important question is how the mean rating responses are expected to differ as a function of the different stimulus levels along the speech continuum. Consider the speech continuum of five levels as illustrated in Figure 8.8. Although perception is categorical, a stimulus is more likely to produce the percept *a* to the extent that it is away from the category boundary and toward the *A* end of the continuum. Variation in the category boundary, the perceptual system, or both allows the percept to have only a probabilistic relation to a given stimulus. A given stimulus produces the percept *a* with probability *Pa* and the percept *b* with probability 1 − *Pa*. Hence, the distribution of rating responses to a given stimulus will actually be a mixture of ratings generated by the two percepts. The proportion of ratings generated from the percept *a* will increase with increases in *Pa*, the likelihood of the percept *a*. Similarly, the proportion of ratings generated from the percept *b* will decrease with increases in *Pa*. The arrows in Figure 8.8 give the mean rating responses resulting from the mixture of the two distributions over trials. Figure 8.9 gives the mean

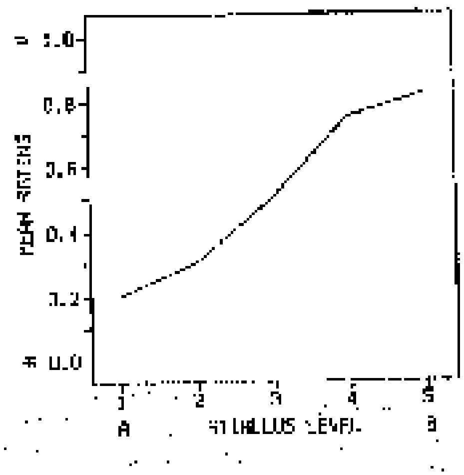

Figure 8.9. Hypothetical-mean-rating responses predicted by both the categorical and continuous models based on the representations in Figures 8.8 and 8.10.

ratings based on the analysis in Figure 8.8 and illustrates that continuous changes in the mean rating response with changes in stimulus level can be predicted by the categorical model.

Continuous model

The continuous model is illustrated in Figure 8.10. In this model, the rating given to a stimulus is a direct function of the percept generated by that stimulus. This model is similar to Thurstone's (1927) law of comparative judgment in which each stimulus is seen as giving rise to a normal distribution of percepts along an internal dimension. In the continuous model, the percepts of two adjacent stimuli will usually differ from each other and the rating responses will reflect this fact (as described in the section entitled "A new test"). In terms of the model in Figure 8.5, the value of *x* given by the sensory stage will differ for each level along the stimulus continuum. The percept of a stimulus toward the *A* end of the continuum will be more *A*-like than that of a neighbouring stimulus towards the *B* side of the continuum. Random variability in the perceptual, memory, or response systems will also result in a distribution of rating responses to any given stimulus. Figure 8.9 shows how the continuous model predicts a systematic and continuous change in mean rating responses with changes in stimulus level. The continuous model, therefore, makes the same predictions as the categorical model with respect to the mean rating responses.

Although the categorical and continuous perception models make similar predictions about [illegible] rating judgments, they are not capable of distinguishing [illegible]. The models can, however, be distinguished on the basis [illegible] the distribution of rating res- [illegible] illustrate the

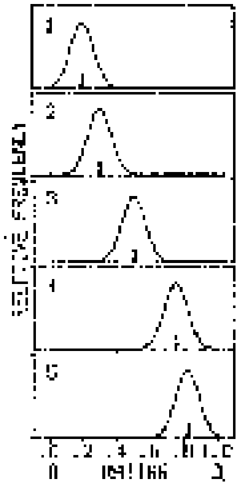

Figure 8.10. Illustration of the continuous-perception-rating model. The distributions depict the likelihood of a percept given each of the five levels (1 through 5) along a hypothetical-stimulus continuum. The location and range of each distribution gives the rating responses that will be observed. The downward-pointing arrows give the mean ratings across the five levels of the stimulus continuum.

overall form of the predicted distribution of rating responses for each of the two models. As can be seen in the figures, although the average rating function (indicated by the arrows) is identical for the two models, the distribution of rating responses is not. For example, at level 3 on the stimulus continuum, the continuous model would predict a single, central distribution, while the categorical model would predict a bimodal distribution with a central trough.

Experimental test

To gather data for the model tests, Massaro and Cohen (1983a) had three groups of subjects rate continua of synthetic-speech stimuli. One group rated consonants differing in place of articulation from /bae/ (as in *bat*) to /dae/, a second group rated consonants differing in voicing from /bae/ to /pae/, and a third group rated vowels on a continuum from /i/ (as in *heat*) to /I/ (as in *hit*). In all cases, the subject rated the degree to which each stimulus represented one category or the other on a continuous scale between the two categories. Because the three continua gave similar results, I will discuss only the voicing data. The stimulus continuum consisted of seven steps, with voice-onset times varying from 0 to 60 msec in steps of 10 msec. As expected from both models, the mean ratings changed relatively continuously as a function of the stimulus level.

The critical feature of our analysis is not the examination of the mean ratings but rather the exact nature of their distribution of occurrence. In order to determine whether the observed distributions of ratings were best fitted by the continuous or the categorical models, the computer program STEPIT (1969) was used. A model is represented to the analysis program STEPIT as a set of prediction equations that

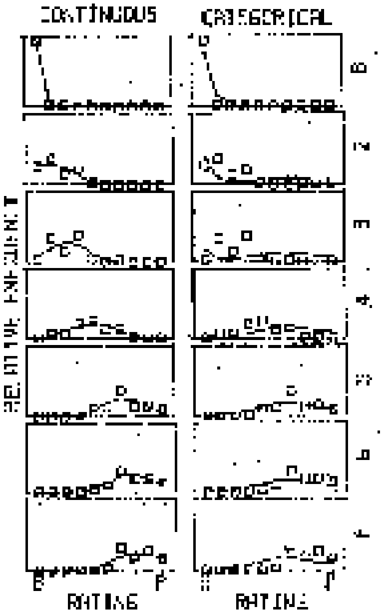

Figure 8.14. Observed and predicted distributions of rating responses for a typical subject as a function of the seven levels along a /bae/-/pae/ stimulus continuum. (After Massaro & Cohen, 1983a.)

contain a set of unknown parameters. By iteratively adjusting the parameters of the model, STEPIT minimizes the chi-square deviations between the observed and predicted values. Thus, what STEPIT does is to find a set of parameter values that, when put in a model, come closest to describing the observed data. We can discriminate between competing models on the basis of the overall goodness of fit. The categorical model required 11 parameters, which included two means and two standard deviations for the two normal distributions and seven sampling probabilities for the seven stimulus levels. For the continuous model there were 14 parameters, which included seven means and seven standard deviations for the seven normal distributions corresponding to the seven stimuli.

Figure 8.14 gives the results for a typical subject, along with the predictions of the continuous and categorical models and demonstrates that the continuous model does a better job of fitting the observed data. What is most noticeable in the figure is how much better the continuous model does for the intermediate stimulus level. Its superiority stems from the observed data appearing to result from a single, central distribution rather than from a mixture of two distinct distributions. The figure also illustrates how much better the continuous model does for the other levels of the distributions. The continuous model much more accurately captures and reproduces the different numbers of observations occurring in the tails of the endpoint stimuli.

The continuous model also gave a better description of the results when contrasted with a categorical model that included a guessing state. This model is the same as the categorical model, but with the additional feature that for each level of the stimulus there is a certain probability that the response is based on a neutral "guessing" distribution. One way of viewing the guessing model would be in terms of a threshold or criterion for making a particular response. If the evidence for one or the other alternative does not exceed the criterion level, then a guessing response is made. Naturally, it would be expected that the guessing response would occur most frequently for the more central, ambiguous levels of the stimulus dimension. This model has 20 parameters: three means and three standard deviations for the /bae/, /pae/, and the guessing distribution, and two sampling probabilities for each of the seven levels along the continuum. Each of the seven stimulus levels needs only two probability parameters to compute because the three probabilities must sum to one. The continuous model was the better descriptor of the results, even with the disadvantage of having six fewer free parameters than the guessing model.

Summary

Building on the developments in psychophysics, it is important to distinguish between sensory and decision stages of processing. Categorical and continuous models of perception made different predictions about the output of the sensory stages. The two theories make different predictions about the distribution of repeated rating judgments to a stimulus along some speech continuum. The results of such a study (Massaro and Cohen, 1983a) provide conclusive evidence that there is continuous information available to the listener in speech perception. The important question, therefore, is how the continuous information is evaluated and integrated in speech perception. In the next section, a fuzzy logical model of perception is developed to provide a comprehensive account of perceptual recognition of speech based on continuous information available to the listener.

Fuzzy logical model of perception

The fuzzy logical model of perception (FLMP) allows a quantitative formulation of the view of continuous-speech perception. Faced with continuous sources of information, the perceiver must evaluate and integrate the sources and then categorize the input. To carry out these functions, the sources must be assigned values that can be easily compared to one another. To achieve this common metric, it is assumed that each of the sources is assigned a fuzzy-truth value based on the degree to which the source supports a given category. It is, therefore, necessary to discuss the concept of fuzzy logic before describing the operations assumed by the model.

Fuzzy logic

In nature it is often the case that things are neither entirely true nor false but rather take on continuous-truth values. For example, we might say that a president is doing a "very bad" job or that a meal is not too spicy. Ordinary logical quantification would require that the president be performing well or not and that the meal be either spicy or not. Fuzzy-logic theory (Goguen, 1969; Zadeh, 1965), on the other hand, allows us to represent the continuous nature of things. In fuzzy logic we can construct a membership function, for example, *short(x)*, which is true to the extent that item *x* is a member of the set *short*. It should be noted that fuzzy truth is different from probability. If we say that a whale is a fish to degree 0.2, this does not mean that there is a 0.2 probability that a particular whale is a fish. Rather, it is true that the whale is a fish to degree 0.2.

An important part of fuzzy logical theory concerns the realization of the standard logical operations of conjunction, negation, and disjunction. The range of truth values, $t(x)$, in fuzzy logic goes from 0 for perfectly false to 1 for perfectly true. Thus, a reasonable definition for negation is the additive complement:

$$t(\bar{x}) = 1 - t(x),$$

where $t(\bar{x})$ is the truth of *not x*. Goguen (1969) has suggested two possibilities for the conjunction of two events, *a* and *b*:

$$t(a \wedge b) = t(a) * t(b),$$

and

$$t(a \wedge b) = \min[t(a), t(b)].$$

With the help of DeMorgan's Law we can derive the two corresponding disjunction operations:

$$t(a \vee b) = t(a) + t(b) - t(a) * t(b)$$

and

$$t(a \vee b) = \max[t(a), t(b)].$$

We should note that either of these definitions reduces to ordinary logic if we restrict truth values to 0 and 1.

Fuzzy logic has been successful in describing logical conjunction and disjunction. In a typical experiment, Oden (1977) investigated which set of definitions of fuzzy logical conjunction and disjunction best fit subjects' judgments about logical combinations of pairs of statements about class-membership functions (e.g., a bat is a bird and a refrigerator is furniture). A large set of such compounds was created by factorially varying the truthfulness of each component. The data from the experiment were best explained by the multiplicative-based rules proposed by Goguen (1969). This does not mean that humans actually carry out the process of multiplication, just that multiplication closely represents the processes involved in information integration (Lopes, 1981; Rumelhart & Norman, 1983). (For further discussion of some of the issues of fuzzy logical representation, see Oden, 1979; Smith and Osherson, 1984, and Massaro, in press.)

Feature evaluation

According to the FLMP (Massaro & Oden, 1980a, b), recognition is carried out in [illegible] during which the information is

transduced by the sensory systems and various features are derived. The features are assumed to be independent of one another and provide continuous information about the degree to which each feature is represented in the speech sound. The outcome of feature evaluation is a truth value, $t(x)$, associated with the presence of each relevant feature.

Feature integration

The second stage of recognition involves the integration of the feature information from the independent sources. During this stage the feature information is compared with perceptual-unit definitions, or prototypes, to determine to what degree each prototype is realized in the speech sound. Prototypes define a perceptual unit in terms of arbitrarily complex fuzzy logical propositions. The outcome of feature integration and prototype matching determines to what degree the stimulus pattern matches each of the candidate prototypes in memory. An important consequence of this process in the model is that one feature has its greatest effect when the second is at its most ambiguous level. Thus, the most informative one has the greatest impact on the judgments.

Pattern classification

The third stage is pattern classification. During this stage the merit of each potential prototype is evaluated relative to the summed merits of the other potential prototypes. The relative goodness of a perceptual unit gives the proportion of its selection or its judged magnitude. This is similar to Luce's (1959) choice rule. In pandemonium-like terms we might say that it is not how loud some demon is calling that counts, but rather the relative loudness of that demon in the relevant crowd of demons. (See also Remez, Chapter 6, and Harnad, Chapter 14, this volume.)

Another new test

It might be asked why another new test is needed given our rejection of the categorical model in favor of a continuous model. There are many reasons. First, no test should be limited to a single domain, and therefore it is important to address other tasks in addition to the rating task. Second, the current test involves identification performance, which is central to speech perception and categorization. Third, the test will allow the formalization and testing of what might be considered to be more general models of categorical and continuous perception. As Conant (1947, p. 36) states, "A theory is only overthrown by a better theory, never by contradictory facts," and the present test provides both fact and theory.

Very few identification experiments have explicitly contrasted formal models of categorical and continuous perception because application of the categorical model has been limited primarily to predicting the relationship between identification and discrimination. In addition, until the Oden and Massaro (1978) model (see also Massaro & Cohen, 1976, 1977), a formal model of continuous perception had not been tested in the context of speech-perception experiments. Another reason has been methodological, as Massaro (1978, 1979) has pointed out. Almost all speech-perception experiments assessed the psychophysical relationship between a single stimulus property and categorization. This functional relationship is not diagnostic of the question of categorical versus continuous perception and does not have the power to test formal models of categorization performance.

Massaro and Cohen (1983b) used the methods of information-integration theory (Anderson, 1981, 1982) to test between categorical and continuous models of speech perception. The primary feature of this approach is to manipulate two or more properties of speech events in a factorial design. The results of such an experiment are much more informative than single-factor experiments. The experiment independently manipulated the audible and visible properties of speech syllables perceivers used in an identification task (see also Chapter 12 by Kuhl, this volume). Using videotape and a speech synthesizer, the authors were able to vary independently the visual and auditory properties of the syllables /ba/ and /da/. The visual property corresponded to the type of visual articulation that was present (in an image of a mouth making a sound). The auditory property corresponded to the values of the second- and third-formant transitions of the syllables. The design involved a visual /ba/, visual /da/, or no visual articulation crossed with nine synthesized speech sounds equally spaced along a /ba/–/da/ continuum. Subjects watched and listened to videotaped presentations of the 27 aural–visual events and identified the event presented on each trial as /ba/ or /da/.

With the single modification of increasing a single-factor to a two-factor design, it is now possible to contrast categorical and continuous models of performance in the task. Following the logic of previous views of CP, it is reasonable to assume that each dimension of the speech event is categorically perceived. There have been many conclusions that the auditory-place continuum between /b/ and /d/ is perceived categorically (Eimas, 1963; Liberman et al. 1957). A similar logic should apply to the visual information (MacDonald & McGurk, 1978). According to a categorical model based on this logic, the listener has only categorical information representing the auditory and visual dimensions of the speech event. This model implies that separate categorical (discrete) decisions are made to the auditory and visual sources (MacDonald and McGurk, 1978). [illegible] identified the purely visual syllables as /ba/ or /da/. According to the formalization of the categorical model, [illegible] the two modalities would be [illegible] independently of one another. That is, separate /da/ or /ba/ decisions would be made to both the auditory and visual sources. The identification of the bimodal syllable would necessarily be based on these separate decisions. Given categorical information from each dimension, there are only four possible outcomes for a particular combination of auditory and visual information: /da/–/da/, /da/–/ba/, /ba/–/da/, or /ba/–/ba/. If the two covert categorizations of the auditory and visual dimensions of a given speech syllable agree, the identification response given the bimodal syllable can follow either source. If the two decisions disagree, it is reasonable to assume that the subject will respond with the decision of the auditory source on some proportion p of the trials, and with the decision of the visual source on the remainder $(1 - p)$ of the trials. In this conceptualization, the magnitude of p relative to $(1 - p)$ reflects the relative dominance of the auditory source.

The probability of a /da/ identification response, $P(D)$, given a particular auditory/visual speech event, A_iV_j, would be:

$$\begin{aligned} P(D : A_iV_j) &= \{1a_iv_j\} + \{pa_i(1 - v_j)\} \\ &\quad + \{(1 - p)(1 - a_i)v_j\} + \{0(1 - a_i)(1 - v_j)\} \\ &= pa_i + (1 - p)v_j, \end{aligned}$$

where i and j index the levels of the auditory and visual stimuli, respectively. The a_i value represents the probability of a /da/ decision, given the auditory level, i, and v_j is the probability of a /da/ decision, given the visual level, j. Each of the four terms in the equation represents the likelihood of one of the four possible outcomes of the separate decisions multiplied by the probability of a /da/ identification response, given that outcome. In the experiment, nine auditory levels are factorially combined with three visual levels. In this model, each unique level of the auditory stimulus would require a unique parameter a_i, and analogously for v_j. Because p reflects a decision variable, its value also requires a unique parameter that would be constant across all stimulus conditions. Thus, a total of $9 + 3 + 1 = 13$ parameters must be estimated for the 27 independent conditions.

Continuous model

The continuous model will be formulated in terms of the FLMP. Applying the model to the present task, using auditory and visual speech, both sources are assumed to provide independent evidence for the alternatives /ba/ and /da/. Defining the important auditory cue as the onsets of the second and third formants and the important visual cue as the degree of initial opening of the lips, the prototypes are

/da/ : slightly falling $F2$–$F3$ & open lips,

/ba/ : rising $F2$–$F3$ & closed lips,

where $F2$–$F3$ represent the onsets of the second and third formants. Given a prototype's independent specifications for the auditory and visual sources, the value of one source cannot change the value of the other source at the prototype-matching stage. In addition, the negation of a feature is defined as the additive complement. That is, we can represent rising $F2$–$F3$ as (1 − slightly falling $F2$–$F3$) and closed lips as (1 − open lips),

/da/ : slightly falling $F2$–$F3$ & open lips

/ba/ : (1 − slightly falling $F2$–$F3$) & (1 − open lips), The integration of the features defining each prototype is evaluated according to the product of the feature values. If a_i represents the degree to which the auditory stimulus, A_i, has slightly falling F2–F3 and v_j represents the degree to which the visual stimulus, V_j, has open lips, the outcome of prototype matching would be:

$$/da/ : a_i v_j$$

$$/ba/ : (1 - a_i)(1 - v_j).$$

If these two prototypes are the only valid response alternatives, the pattern classification operation would determine their relative merit leading to the prediction that

$$P(D : A_i V_j) = \frac{a_i v_j}{a_i v_j + [(1 - a_i)(1 - v_j)]}$$

Given nine levels of A_i and three levels of V_j in the present task, the predictions of the model require 9 + 3 = 12 parameters, one fewer than the categorical model.

Experimental test

Each point in Figure 8.12 represents the proportion of /da/ identifications as a function of the auditory level; the curve parameter is the visual condition. Both the auditory and visual sources influenced identification, with the contribution of the visual source larger at the middle range of the auditory continuum. The lines in the left panel of Figure 8.12 give the average predictions of the continuous model applied to the individual results of each of seven subjects. The right panel gives the predictions for the categorical model. As the figure shows, the continuous model gave a significantly better account of the identification judgments with a RMSD of less than half the RMSD of the categorical model (Massaro & Cohen, 1983b).

Identification-reaction times

Reaction times of identification responses might also be used to assess continuous and categorical models of speech perception. The most direct interpretation of the CP assumption is that the reaction times of a given identification response should not vary with changes in the speech stimulus. If subjects cannot discriminate differences within a category then there is no basis for differences in reaction times to the different stimuli. A continuous-perception model predicts that the reaction time for identification would depend on the degree to which the sensory information provided unambiguous support for a given category (Norman & Wickelgren, 1969). If different within-category stimuli required different amounts of time for identification then these time differences would provide a source of information for discriminating within-category stimuli. Previous studies have revealed increases in identi- [illegible]

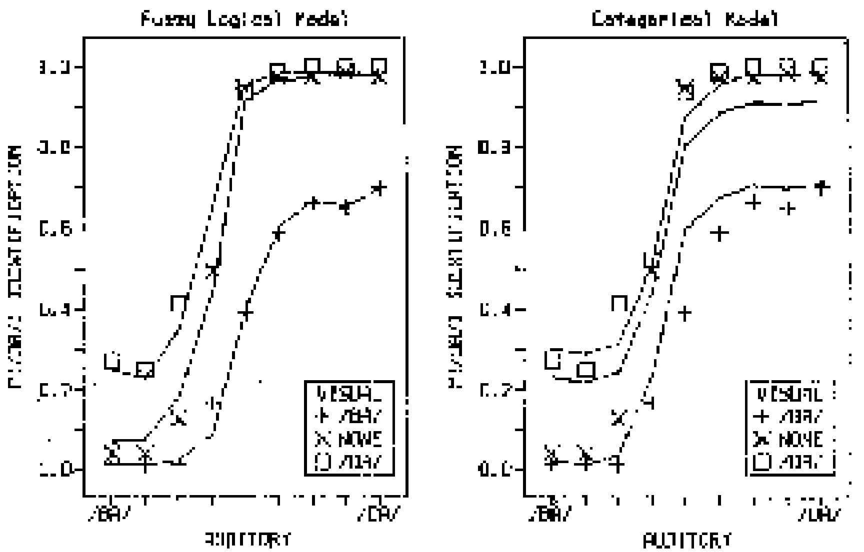

Figure 8.12. Proportion of observed (points) and predicted (lines) /da/ identifications as a function of the auditory and visual levels of the speech event. (After Massaro & Cohen, 1983b.) The predictions of the continuous fuzzy logical model are given in the left panel and the predictions of the categorical model are given in the right panel.

identification process. Analogous to the study of information processing in other domains, reaction times have the potential for making perceptual recognition processes transparent and for addressing the issue of categorical versus continuous speech perception.

Experimental test

We have evaluated reaction times to assess whether a source of information is continuous or categorical in feature evaluation. By independently varying two sources of information in the visual/auditory experiment, it was possible to provide quantitative tests of the continuous and categorical views of feature evaluation. For the analysis, the reaction times of the identification judgments were pooled across /ba/ and /da/ responses before the average reaction times were computed. Figure 8.13 gives the average identification-reaction times as a function of the auditory and visual levels of the speech event. The reaction times varied significantly with the auditory level, decreasing about 100 msec as the auditory level went from /ba/ to /da/, probably because the auditory continuum was biased toward the sound /da/. The overall response probability for /da/ was higher than for /ba/. Reaction times are known to decrease with increases in response probability (Sternberg, 1969, [illegible]

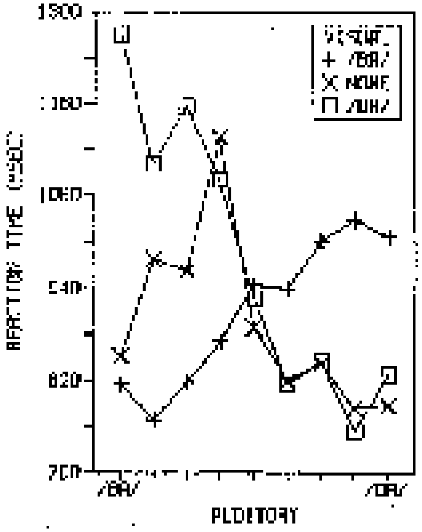

Figure 8.13. Identification-reaction times as a function of the auditory and visual levels of a speech event. (After Massaro & Cohen, 1983b.)

The significant interaction shown in Figure 8.13 shows that reaction times given a visual /ba/ articulation tend to increase with changes from /ba/ to /da/ along the auditory continuum. An analogous result occurred for the visual /da/ in that reaction times tend to decrease with auditory changes from /ba/ to /da/. There are two possible explanations for these results. First, subjects were significantly faster when the auditory and visual information agreed than when they conflicted. Second, the reaction time is positively related to the ambiguity of the speech event.

Ambiguity versus conflict

Although these two explanations make similar predictions for some of the speech events, conflicting auditory and visual information is not always correlated with the ambiguity of a speech event. Ambiguity can be defined in terms of the degree to which the probability of a /da/ identification differs from 0.5, and a given probability of a /da/ identification could result from a variety of levels of conflicting information. A very /da/-like sound paired with a good /ba/ articulation would give the same probability of /da/ identification as a very ambiguous sound paired with a relatively ambiguous /da/ articulation. The first case would have greater conflicting information than the second, although both would be characterized by the same amount of ambiguity.

The fuzzy logical model of perception is consistent with the result that reaction times increase with increases in ambiguity of the speech event. To the extent that listeners have ambiguous information, it should take longer to decide on one of the discrete-response alternatives at the pattern-classification stage (Thomas & Myers, 1972; Norman & Wickelgren, 1969; Smith, 1968). On the other hand, the increase in reaction times with increases in conflict of the visual and auditory sources is not necessarily consistent with the fuzzy logical model. Given independent sources of information and a multiplicative integration, the time for feature evaluation and prototype matching should not necessarily change with the degree of visual and auditory conflict. Accordingly, determining which explanation best accounts for the reaction times provides a new test of the fuzzy logical model of perception.

The categorical model has no mechanism to account for an effect of ambiguity, but would seem to predict some effect of conflict. We might expect that different decisions to the auditory and visual sources would lead to longer identification times than the same decision to the two sources. Accordingly, conflict should be related to the likelihood that the two sources give different decisions. To quantify the idea of conflicting information, we used the parameter values for the auditory and visual sources of information given by the fuzzy logical model. The degree of conflict was taken to be the absolute difference between the parameter values specifying the auditory and visual sources of information:

$$C_{ij} = |\text{visual } D_i - \text{auditory } D_j|,$$

where C_{ij} is the conflict given by the ith level of visual articulation and the jth level along the auditory continuum.

Ambiguity A was defined as the absolute difference between $P(/da/)$, the probability of a /da/ identification, and 0.5

$$A_{ij} = |P(/da/) - 0.5|.$$

In this conceptualization, the lowest ambiguity value possible would be 0.5 while the highest would be 0.

Results

To evaluate the conflict and ambiguity explanations, a critical result is the performance on neutral trials in which no articulation is given. In this case, the ambiguity of the auditory source of information will vary without a concomitant variation in the conflict between it and the visual source. If reaction times increase with increases in ambiguity along the auditory dimension, the result cannot be accounted for by conflict. In fact, the reaction times to neutral trials increased toward the middle of the auditory continuum. This result replicates previous studies that evaluated identification times as a function of only a single source of information (Pisoni & Tash, 1974; Repp, 1981; Studdert-Kennedy, Liberman, & Stevens, 1963). Accordingly, reaction times increased with increases in the ambiguity of the auditory stimulus even though there was no change in conflict between the auditory and visual sources.

For a complete comparison of the conflict explanation with the ambiguity explanation, correlations between conflict and ambiguity with reaction time were carried out on the results of individual subjects in the study by Massaro and Cohen (1983b, Experiment 2). Multiple-regression analyses across the seven subjects revealed that ambiguity accounted for 26% of the reaction time variance that could not be accounted for by conflict. Conflict accounted for only 3% of the variance that could not be accounted for by ambiguity. This result seems to resolve the question in favor of continuous perception of the auditory and visual sources of information, with identification time a direct function of the ambiguity of the speech event.

Conclusion

I have tried to refute the theory of CP, drawing on a broad range of methodological, theoretical, and experimental issues. At the methodological level, we have learned that the identification/discrimination task provides no support for CP. First, the categorical model usually provides an inadequate description of the results, and it has not been shown to provide a better description than alternative models. Second, even if the results in the task provided unequivocal support for the categorical model, explanations other than CP are possible.

At the theoretical level, it is necessary to distinguish between sensory and decision processes in the categorization task. What is central for our purposes is that decision processes can transform continuous sensory information into results usually taken to reflect CP. Thus, finding relatively categorical partitioning of a set of stimuli in no manner implies that these stimuli were perceived categorically.

The most convincing evidence against CP comes from direct experimental tests between categorical and continuous models of perception. Three new results were brought to bear on the issue of categorical versus continuous perception. First, subjects were asked to make repeated ratings of the degree to which a stimulus represents a given category. The distribution of the rating judgments to a given stimulus is more adequately described by a continuous than a categorical model of perception. Second, subjects asked to classify speech events independently varying along two dimensions produce identification results consistent with the assumption of continuous information along each of the two dimensions. A model based on categorical information along each dimension gives a very poor description of the identification judgments. Third, the reaction times of identification judgments illustrate that members within a speech category vary in ambiguity or the degree to which they represent the category. The best conclusion is to reject all reference to categorical perception of speech and to concentrate instead on the structures and processes responsible for categorizing the world of speech.

[illegible] New York: Academic Press.
[illegible] University Press.

Barclay, J. R. (1972). Noncategorical perception of a voiced stop: A replication. *Perception & Psychophysics, 11*, 269–273.

Conant, J. B. (1947). *On understanding science*. New Haven, CT: Yale University Press.

Cutting, J. E. (1982). Plucks and bows are categorically perceived, sometimes. *Perception & Psychophysics, 31*, 462–476.

Cutting, J. E., & Rosner, B. S. (1974). Categories and boundaries in speech and music. *Perception & Psychophysics, 16*, 564–570.

Eimas, P. D. (1963). The relation between identification and discrimination along speech and non-speech continua. *Language and Speech, 6*, 206–217.

Elman, J. L. (1979). Perceptual origins of the phoneme boundary effect and selective adaptation to speech: A signal detection theory analysis. *Journal of the Acoustical Society of America, 65*, 190–207.

Fujisaki, H., & Kawashima, T. (1969). On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo, 28*, 67–73.

Goguen, J. A. (1969). The logic of inexact concepts. *Synthese, 19*, 325–373.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hary, J. M., & Massaro, D. W. (1982). Categorical results do not imply categorical perception. *Perception & Psychophysics, 32*, 409–418.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431–461.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358–368.

Lopes, L. (1981). Decision making in the short run. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 377–385.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics, 24*, 253–257.

Massaro, D. W. (1975a). *Experimental psychology and information processing*. Chicago: Rand-McNally.

Massaro, D. W. (Ed.) (1975b). *Understanding language: An information processing analysis of speech perception, reading and psycholinguistics*. New York: Academic Press.

Massaro, D. W. (1976). Auditory information processing. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4): *Attention and memory*. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1978). A stage model of reading and listening. *Visible Language, 12*, 3–26.

Massaro, D. W. (1979). Reading and listening (Tutorial paper). In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1, pp. 331–354). New York: Plenum Press.

Massaro, D. W. (1984). Building and testing models of reading processes: Examples from word recognition. In P. D. Pearson (Ed.), *Handbook of reading research*. New York: Longman.

Massaro, D. W. (in press). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

Massaro, D. W., & Cohen, M. M. (1976). The contribution of fundamental frequency and voice onset time to the /zi/–/si/ distinction. *Journal of the Acoustical Society of America, 60*, 704–717.

Massaro, D. W., & Cohen, M. M. (1977). The contribution of voice-onset time and fundamental frequency as cues to the /zi/–/si/ distinction. *Perception & Psychophysics, 22*, 373–382.

Massaro, D. W., & Cohen, M. M. (1983a). Categorical or continuous speech perception: A new test. *Speech Communication, 2*, 15–35.

Massaro, D. W., & Cohen, M. M. (1983b). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 753–771.

Massaro, D. W., & Oden, G. C. (1980a). Evaluation and integration of acoustic features in speech perception. *Journal of the Acoustical Society of America, 67*, 996–1013.

Massaro, D. W., & Oden, G. C. (1980b). Speech perception: A framework for research and theory. In

N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 3). New York: Academic Press.

McNabb, S. D. (1976). Must the output of the phonetic feature detector be binary? *Research in Speech Perception Progress Report* (Vol. 2). Bloomington, IN: Department of Psychology, Indiana University.

Norman, D. A., & Wickelgren, W. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, 6, 192–208.

Oden, G. C. (1977). Integration of fuzzy logical information. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 565–575.

Oden, G. C. (1979). A fuzzy logical model of letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 336–352.

Oden, G. C. (1981). Fuzzy propositional model of concept structure and use: A case study in object identification. In G. W. Lasker (Ed.), *Applied systems research and cybernetics*. Elmsford, NY: Pergamon Press.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, 85, 172–191.

Paap, K. R. (1975). Theories of speech perception. In D. W. Massaro (Ed.), *Understanding language*. New York: Academic Press.

Pastore, R. E., Ahroon, W. A., Baffuto, K. J., Friedman, C., Puleo, J. S., & Fink, E. A. (1977). Common-factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 686–695.

Pisoni, D. B. (1973). Auditory and phonetic codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13, 253–260.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 297–314.

Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America*, 55, 328–333.

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, 15, 285–290.

Repp, B. H. (1981). Perceptual equivalence of two kinds of ambiguous speech stimuli. *Bulletin of the Psychonomic Society*, 18, 12–14.

Repp, B. H. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 10). New York: Academic Press.

Rosen, S. M., & Howell, P. (1981). Plucks and bows are not categorically perceived. *Perception & Psychophysics*, 30, 156–168.

Rumelhart, D. E., & Norman, D. A. (1985). Representation in memory. *Center for Human Information Processing*, 116, 1–117.

Sawusch, J. R. (1977). Peripheral and central processes in selective adaptation of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 62, 738–750.

Siegel, J. A., & Siegel, W. (1977a). Absolute identification of notes and intervals by musicians. *Perception & Psychophysics*, 21, 143–152.

Siegel, J. A., & Siegel, W. (1977b). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics*, 21, 399–407.

Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337–361.

Smith, P. T. (1974). Cross-discriminability and response bias. *British Journal of Mathematical and Statistical Psychology*, 27, 35–48.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358–1368.

Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., & Cooper, F. S. (1970). The motor theory of speech perception: A reply to Lane's critical review. *Psychological Review*, 77, 234–249.

Studdert-Kennedy, M., Liberman, A. M., & Stevens, K. N. (1963). Reaction times to synthetic stops and vowels at phoneme centers and at phoneme boundaries. *Journal of the Acoustical Society of America*, 35, 1900.

Theios, J., & Walter, D. G. (1974). Stimulus and response frequency and sequential effects in memory scanning reaction times. *Journal of Experimental Psychology*, 102, 1092–1099.

Thomas, E. A. C., & Myers, J. L. (1972). Implications of latency data for threshold and nonthreshold models of signal detection. *Journal of Mathematical Psychology*, 9, 253–285.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34, 273–286.

Watson, C. S., Rilling, M. E., & Bourbon, W. T. (1964). Receiver-operating characteristics determined by a mechanical analog to the rating scale. *Journal of the Acoustical Society of America*, 36, 283–288.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.