IS THERE A NATURAL SENSITIVITY AT 20 MS IN RELATIVE TONE-ONSET-TIME CONTINUA? A REANALYSIS OF HIRSH'S (1959) DATA*

Stuart Rosen
Department of Phonetics and Linguistics
University College London
4 Stephenson Way,
London, NW1 2HE, U.K.

Peter Howell
Department of Psychology
University College London
Gower St.
London, WC1E 6BT, U.K.

Over the last dozen years, there has been much speculation concerning the possibility of auditory underpinnings to the structure of phonemic categories. Support for this hypothesis is usually drawn from studies of categorical perception, a mode of perception in which it was originally supposed that sounds can only be discriminated from one another to the extent to which they are labelled differently. This is meant to contrast with the more common situation, often called "continuous perception", in which the ability to discriminate far outstrips the ability to label differentially (Miller, 1956). Practically speaking, a continuum is said to be categorically perceived when there is: (1) a sharp categorization function, (2) a peak in the discrimination function at the category boundary and (3) near-chance discrimination performance within categories (Studdert-Kennedy, Liberman, Harris, and Cooper, 1970). Categorical perception was initially thought to be confined to speech sounds, and to arise from a reference to some aspect of the articulatory process (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). Since then, the phenomena of categorical perception have been obtained with a number of nonspeech auditory continua, and even with visual stimuli (e.g., Pastore, Ahroon, Baffuto, Friedman, Puleo, and Fink, 1977). These studies gave rise to the idea that "categorical perception" need not rely on the use of categories at all, but could arise simply from a non-uniform discriminability across the stimulus continuum. Few would argue against such a possibility. More controversial are the notions, put forward most concisely by Stevens (1981), that these auditory sensitivities are responsible for the categorical perception of speech sounds, and are the basis for all phonemic categories.

There is relatively little evidence supporting this point of view. Restricting ourselves only to studies with adult human listeners, all of these concern the voicing distinction in initial homorganic plosive stops[1]. Both Miller, Wier, Pastore, Kelly, and Dooling (1976) and Pisoni (1977) have demonstrated categorical perception of nonspeech continua that have acoustic characteristics reminiscent of those which can signal the voicing distinction in plosives. Pisoni's work has attracted rather more scrutiny because he claimed that the acoustic contrast incorporated in his continuum is the primary acoustic contrast employed by listeners in perceiving the voice-onset-time (VOT)

*This work has been supported by the Medical Research Council of the United Kingdom.

contrast. We will not address that issue here. Our current concern will be Pisoni's finding of a natural auditory sensitivity at a, perhaps, not intuitively predictable point. Even more oddly, it is widely supposed that this result is consonant with what Hirsh (1959) found in a set of similar, although more wide-ranging experiments. In fact, Hirsh's results are quite different from Pisoni's in a number of ways. We shall see that Hirsh's study has been misinterpreted since its first appearance and, ironically, has been marshalled in support of both sides of a controversy concerning the adequacy of auditory processes as a basis for the categorical perception of VOT continua.

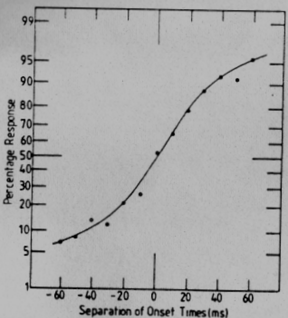

FIGURE 1, Mean results from a task requiring the judgement of the relative order of onset of two tones. Points to the right of 0 on the abscissa indicate that the higher tone led the lower. Points to the left of 0 indicate that the lower tone led the higher. The ordinate indicates the percentage of times the subject reported that the high tone led the low. The smooth curve is fit by eye to the original data points. Redrawn from Figure 4 of Hirsh (1959) including only those points that deal with two-tone stimulus complexes.

Hirsh's (1959) primary interest was in the ability of listeners to determine the order of occurrence of two sounds as a function of the temporal disparity between their onsets. A variety of different sounds were used, but we need only concern ourselves with the results with two pure tones, as these are most comparable to Pisoni's (1977) stimuli. The experimental paradigm was simple. A particular stimulus consisted of two sinusoids of different frequencies, with a range of relative onset times varying from -60 to +60 ms (i.e., the low frequency tone starting 60 ms before the high frequency tone to the high starting before the low by 60 ms). The two sinusoids terminated simultaneously and were about 500 ms long. Five subjects listened to the stimulus complex played repetitively until they could decide which of the two tones came first, the lower or the higher. Figure 1 shows the results, averaged over the five listeners and the five possible frequency pairs.

The accuracy of the subjects' responses improves smoothly with increasing separation of onset times. Hirsh, as is commonly done, summarized this result by choosing the point at which the subjects were performing at a level of 75% correct, concluding that a little less than 20 ms was needed to resolve the order of events. Note also that

there is no constant error: when the onsets of the two stimuli are simultaneous, subjects say the high tone leads the low as often as they say the low leads the high.

The first suggestion that the discriminability of Hirsh's stimuli might be non-uniform was advanced by Miller et al. (1976) who supposed that distinct percepts occur as one increases the separation of the two onsets: "Thus, as the amount by which the high tone precedes the low is increased, perceptual boundaries or thresholds are crossed corresponding to the perceptual effects of nonsimultaneity, Gestalt sequence with obvious ordering, and ordered onsets of two distinct percepts... one would expect to find perturbations in the Weber fraction at the perceptual boundaries." Hirsh did not perform a discrimination experiment but such perturbations should evidence themselves in the labelling function. As noted above, however, Hirsh's data shows a smooth increase in subject performance with increasing onset separation, and no evidence of discontinuities.

Pisoni (1977) took Miller et al.'s suggestion and performed both labelling and discrimination experiments with a continuum closely modeled on Hirsh's stimuli, with some slight differences. Pisoni used a single two-tone complex with components at 500 and 1500 Hz. Their relative onset times varied from -50 to +50 ms. For the labelling experiment, subjects were first trained with feedback to respond appropriately to the endpoint stimuli, -50 and +50 ms. After 320 trials in a random order, the -30 and +30 ms stimuli were introduced for a further 160 trials with feedback. During the identification experiment proper, all 11 stimuli were presented in a random order (15 presentations per stimulus) without feedback. Pisoni presented the results for each of the eight subjects individually. For six of the eight, the category boundary occurred not at 0 ms, as it did for Hirsh's subjects, but in the region where the high tone led the low by 15 to 25 ms. In order to compare the results from the two studies in more detail, we took the average results from both Hirsh (1959) and Pisoni (1977) and fitted cumulative normal curves to them using a maximum-likelihood technique (Bock and Jones, 1968, also known as probit analysis). The estimated category boundary was 0 ms for Hirsh's study and 13.6 for Pisoni's. Also, the slope of the curve fit to Hirsh's results was about 1.7 times shallower than the slope fit to Pisoni's indicating that Hirsh's subjects were rather less sensitive to changes in relative tone-onset-time than Pisoni's were. This is also reflected in the fact that Hirsh's subjects never did better than a 95% correct labelling of the order of the tones, even when there was 60 ms between the two onsets, while Pisoni's subjects averaged about 99% correct for onset asynchronies of +50 ms.

There are a number of differences between Hirsh's and Pisoni's studies which may be responsible for the discrepancy in the obtained category boundary. The three minor differences in the stimuli do not seem to account for the differences in results. Firstly, although Pisoni used other frequencies for his two tones than Hirsh did, Hirsh used a variety of frequency pairs (250-300 Hz, 250-1200 Hz, 250-4800 Hz, 1000-2000 Hz and 1000-4800 Hz) and obtained nearly the same results in all cases. Secondly, Hirsh's tones were typically at nearly the same intensity level[2], whereas Pisoni's 1500 Hz tone was 12 dB lower in intensity than the 500 Hz tone. Pisoni, however, varied the upper frequency tone over a 24 dB range (-12 dB to +12 dB re the level of the lower tone) and found no differences in the category boundaries obtained in labelling. Finally, Hirsh's stimuli had a base duration of 500 ms, whereas Pisoni's were 230 ms. Perhaps category boundaries on tone-onset-time continua shorten with increasing base duration. Pastore, Harris, and Kaplan (1982), however, found category boundaries to increase with increasing stimulus duration.

Having ruled out stimulus differences, then, the most likely cause of the discrepancy is the task set the subjects. Differences in instructions to the subjects may have caused them to attend to different aspects of the stimulus complex. Hirsh asked his to identify which of the two tones came first - instructions that favour a category boundary at simultaneity. Pisoni trained his subjects to respond differentially to exemplars of each category on the continuum and gave no verbal labels to the important stimulus characteristics. One other factor may be important. Pisoni gave feedback in initial training with stimuli at relatively extreme positions on the continuum. Hirsh (1959) makes no explicit statement about whether feedback was used or not, but informs us (Hirsh, personal communication) that feedback was never given. Had feedback been given on each trial, this might have encouraged the subjects to place their boundary at a value near 0 ms in order to maximize their performance.

A few other studies have investigated the labelling of tone-onset-time continua. Pisoni (1980) used the same continuum as previously in an adaptation study. Only the baseline results (i.e., an ordinary labelling experiment) need concern us here. Four such experiments were run. The obtained boundary (averaged over subjects) varied from 8.3 to 17.9 ms, for an overall mean of 13 ms, nearly identical to the value we calculated previously for the 1977 study.

Summerfield (1982) also explored the perception of tone-onset-time continua. Four frequency pairs were possible with 2500 Hz always as the upper frequency tone. The lower frequency tone was at 200, 300, 400, or 500 Hz. The relative levels of the tone were set to match the relative levels of the formants in a set of analogous voice-onset-time stimuli, and so varied with the frequency of the lower tone in the pair. The level of the upper tone was therefore -8, -5, 0 and, +4 dB relative to the level of the lower tone for the frequencies of 200 to 500 Hz, respectively. In Experiment 1, Summerfield used an adaptive technique to estimate the category boundary and found the four-subject mean to decrease monotonically from 20.6 ms to 11.7 ms with increases in the frequency of the lower tone. In a more typical labelling experiment, in which all the stimuli occurred equally often, there was no significant change of the category boundary with frequency of the lower tone. The mean boundary (over subjects and conditions) was 21.5 ms. This is quite different from the 13 or so ms Pisoni has consistently found, and very different from Hirsh's boundary of 0 ms. In this latter experiment however, no negative tone-onset-times occurred, only values between 0 and +60 ms. As is well known (Parducci 1965), the range of values used in labelling experiments is an important determinant of the category boundary obtained for both nonspeech and speech (see Howell and Rosen, 1984, for a review) so it is perhaps not surprising that Summerfield found a longer category boundary in this case. If range effects were important, his own adaptive studies, which used a continuum extending down to values of -20 ms, would be expected to

show shorter category boundaries. Consistent with this prediction, the mean category boundary was 14.9 ms (ignoring the influence of the frequency of the lower tone). It also seems likely that Summerfield's instructions to label the stimuli as to whether the onsets of the two component tones were "simultaneous" or "successive" would encourage subjects to place their category boundary at a positive relative onset-time.

It is important to resolve these serious differences, especially in the light of common misinterpretations of what Hirsh's data actually show. There are many comments in the literature (one of them ours) that imply or state that Hirsh's data show some evidence for a discontinuity in discriminability near 20 ms. This basic misinterpretation has to do with the nature of a psychometric function. Hirsh found performance in his experiments to vary smoothly with increasing onset separations, and, as we pointed out above, followed tradition in choosing the 75% correct point (17 ms) to summarize his data. He could just as well have picked 60%, giving about 5 ms as the crucial time, or, perhaps more appropriately for linguistic use (to ensure more reliable reception), 90% giving about 30 ms. None of these choices is preordained.

The fallacy in reasoning may be more easily appreciated when applied to a simple psychophysical continuum. Suppose subjects were presented with a continuum of sounds varying from 60 to 70 dB SPL in 1 dB steps and asked to label them as "loud" or "soft". We might find they had a category boundary at 65 dB and, by interpolation, that a stimulus of 66.5 dB was necessary for the label "loud" to occur 75% of the time. If we consider all stimuli above 65 dB as "loud", we might also say that presenting a sound at 66.5 dB led to 75% correct performance. So far so good. But would we then want to argue that the pair of stimuli 65 and 67 dB were inherently more discriminable than the pair 68 and 70 dB because they straddled this (arbitrarily defined) 75% point?

By the same token, there is no reason to suppose from Hirsh's labelling function alone that discriminability between tone-onset-time stimuli will be better around 17 ms. On the other hand, we do not preclude such a possibility. This requires a somewhat more complicated analysis.

We can predict a discrimination function from Hirsh's labelling data under some simple and fairly reasonable assumptions. Using Thurstonian Case V analysis (assuming a uni-dimensional psychological continuum where stimulus densities are Gaussian distributed with equal variance; see Torgerson, 1958, for details) we take the normal deviates of the proportion of "high precedes low" judgements (Figure 1) as the scale values for each of the stimuli. In order to predict performance in a discrimination task for, say, two-step comparisons (usually the most informative with stimulus spacings commonly used, as three-step comparisons often show ceiling effects near perfect performance and one-step comparisons often show floor effects at chance), the d' values between the appropriate stimuli are computed by taking the difference between the scale values[3]. We can then convert these values into the proportion correct that might be expected in an ABX task using a method based on signal detection theory developed by Macmillan, Kaplan, and Creelman (1977) and the tables of Kaplan, Macmillan and Creelman (1978). Since this transformation increases monotonically, it will preserve any peaks in the d' function, so for that purpose it is not crucial whether we examine d' or proportion correct: we use the latter for a later comparison to Pisoni's (1977) results. Figure 2 shows the final outcome.

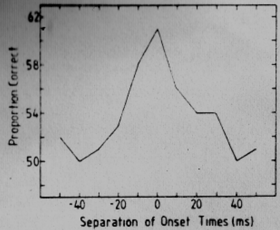

FIGURE 2, ABX discrimination results for stimuli varying in relative tone-onset-time predicted from Hirsh's (1959) data of Figure 1.

As we argued from the labelling function, there is no increased sensitivity around tone-onset-times of 20 ms. On the other hand, discriminability is not uniform across the continuum. The discrimination function is significantly peaked, with best discrimination near simultaneity, 0 ms. Under the assumptions we have made, performance is simply a monotonic transformation of the derivative of the labelling curve. Therefore, the peak is a reflection of the fact that the labelling function is steepest at its centre and flattens towards its edges, as is clearly seen in Figure 1.

This result might be expected if Weber's Law described the discrimination of relative onset time. Remember that we have made predictions for discrimination based on a fixed 20 ms difference between stimuli; if the difference limen for different values of relative onset time was proportional to the magnitude of the relative onset time (as would happen if Weber's Law applied), then the constant 20 ms change should lead to best performance at the shortest relative onset times (i.e. simultaneity), which decreased monotonically with increasing relative onset time, just as we have found. This also implies that the original labelling function of Figure 1 should be more linear under a logarithmic transformation of the relative-onset-time axis. We leave out of consideration 0 ms relative onset time as this value cannot be log-transformed. Taking advantage of the symmetry of Hirsh's results around 0 ms, we average the absolute values of the normal deviates corresponding to performance at each of the six onset times from 10 to 60 ms. (In other words, we average the performance at -10 ms with that of 10 ms, -20 ms with that of 20 ms, and so on.) These values are plotted as a function of linear and log relative onset time in Figure 3.

With the linear scaling on the left of Figure 3, we see the curvature of the function already displayed in Figure 1: a decreasing slope with increasing onset time. When the stimuli are plotted on a logarithmic scale, however, the slope of the curve seems to remain constant out to the longest relative onset times measured. This is further support for the notion that Weber's Law holds for these stimuli, and that no special acuity exists for relative onset times near 20 ms.

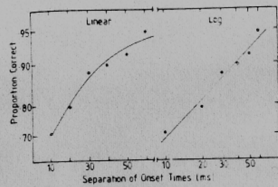

FIGURE 3, Linear and logarithmic scaling of Hirsh's (1959) data (seen in Figure 1), averaged over positive and negative relative onset times, as a function of the absolute value of the relative onset times. The value for a separation of 0 ms is not used because it cannot be logarithmically scaled. The smooth curve on the left is fit by eye to the data points; that on the right is a least-squares straight-line fit.

Pisoni (1977) found quite a different result when he presented his stimuli in a standard ABX two-step (20 ms) paradigm. Although he also found a primary peak in the discrimination function, it occurred at a value of 10 to 20 ms. This correspondence between the peak of the discrimination function and the category boundary obtained in a labelling experiment was taken as evidence that the tone-onset-time continuum was categorically perceived[4]. Pisoni's data is quite convincing on this point, and it was therefore argued that both the labelling boundary and the discrimination peak were due to auditory processes. This claim was strengthened by the demonstration that the subjects' discrimination performance was not a result of their using the previously-learned labels; subjects with no previous experience of identifying the stimuli showed the same pattern of discrimination. Contrary to many claims, however, this result is not compatible with Hirsh's (1959) findings.

Hirsh's results have, however, a long history of misinterpretation by workers in speech perception. They were first cited, soon after they appeared, in a study by Liberman, Harris, Kinney and Lane (1961) who considered Hirsh's stimuli to incorporate a similar acoustic contrast to their own VOT stimuli. Before we can address that issue, however, we will need to describe that study in detail.

Liberman et al. (1961) demonstrated categorical perception for a /do/ to /to/ continuum in which the onset of the first formant (F1) relative to the upper formants was progressively delayed from 0 to 60 ms in 10 ms steps. A nonspeech control condition was also included in which the Pattern Playback schematic spectrograms were turned upside down and further modified before use. These were not heard as speech and resulted in a continuum in which the third and highest formant had its onset delayed relative to the two lower formants. Strictly speaking, it was not possible to say anything about the categoricalness of the perception of the nonspeech stimuli as they were not presented for labelling, only in the ABX discrimination paradigm. Discrimination performance for these stimuli was much inferior to that obtained for the speech stimuli, however, and it was concluded that (assuming the control to be fair) the performance obtained with speech could be interpreted as an example of distinctiveness acquired through learning.

Liberman et al. also compared their results to those obtained by Hirsh, whose stimuli they considered another nonspeech control for the F1-cutback stimuli. Two claims were made: first, that the overall performance exhibited by Hirsh's subjects judging the nonspeech continuum was inferior to Liberman et al.'s subjects judging the speech continuum, and second, that the speech discrimination functions were peaked, unlike the discrimination functions from nonspeech.

Liberman et al.'s primary error is a misinterpretation of the nature of Hirsh's experiment. They seem to consider Hirsh's results as a discrimination function, instead of the labelling function it is. Thus, Hirsh's finding of 75% correct performance with relative onset times of about 17 ms is compared to their own finding of 75% correct performance in an ABX task with less than 12 ms difference in time of onset on the F1-cutback continuum. Only the speech labelling functions of Liberman et al. can be directly compared to Hirsh's results. This comparison, when made, leads to much stronger support, in fact, for Liberman et al.'s assertion of acquired distinctiveness for the speech sounds. They state ".... in every case a change of 10 msec in the first-formant cutback is sufficient to shift the responses from 75% /d/ to 75% /t/..." while Hirsh's data show that a change of about 35 ms in relative onset time is necessary to shift the responses from 75% "low leads high" to 75% "high leads low".

Our analysis summarized in Figure 2 is needed to address Liberman et al.'s second claim: "One finds in Hirsh's results no indication of the sharp peaks so clearly evident in the discrimination functions of the present experiment". This is equivalent to looking for sharp peaks in the speech labelling functions! The discrimination function we have predicted from Hirsh's data does, in fact, show a clear peak.

On the other hand, here we can confirm again Liberman et al.'s assertion that performance with the speech continuum is far superior to that obtained with nonspeech. For a 20 ms difference in relative onset time, we predict that Hirsh's subjects would obtain, at best, 61% correct in an ABX discrimination task. Liberman et al.'s subjects, as noted above, did about 75% correct with differences of slightly less than 12 ms.

This reanalysis of Hirsh's data in no way impugns the substance of what Liberman et al. (1961) were saying. As we have noted, performance with the speech stimuli is much better than that obtained for the nonspeech stimuli. Furthermore, even though Hirsh's data may, contrary to Liberman et al.'s assertions, contain discrimination peaks, they are not in the same place as the peaks for the speech continuum (about 20 to 30 ms).

If however, we take Pisoni's (1977) results as a nonspeech control condition, few of these points hold. Taking only the data from the 5 best subjects of 8 (Liberman et al. used the best 11 of 20), the labelling functions seem to be as sharp as, and their discrimination abilities at least roughly equivalent to Liberman et al.'s subjects. Also, there is a significant peak in the discrimination function at a value not too far from the one obtained with speech.

Many questions remain. First, there is the problem of the discrepancy between Hirsh's (1959) findings and the more recent ones of Pisoni (1977, 1980) and Summerfield (1982). Second, even if these later results are upheld, their theoretical underpinnings are shaky. Hirsh's data seemed explicable on the basis of a limitation in the ability to identify the order of events since the labelling function he obtained was symmetric. Pisoni and Summerfield's asymmetric labelling functions imply there is more to it than that, although Pisoni argues explicitly that it is such a limitation that underlies both the position of the category boundary and the peak in the discrimination function. If this were the case, one would expect similar discrimination peaks in the continuum when the low tone led the high. Pisoni's (1977) discrimination data from untrained subjects (Experiment II), when averaged, do show evidence of a secondary discrimination peak in this region, although it is considerably smaller than the peak in the region where the high tone leads the low. Furthermore, of the 11 subjects who performed at levels above chance (one subject did not), only 6 show convincing discrimination peaks in the "low leads high" region, while all show peaks in the "high leads low" region.

Third, there is the problem of the variability of the obtained category boundaries even in the modern studies. If the labelling function is determined by a natural auditory property, the obtained boundaries should be invariant over manipulations in the range of stimuli presented. Our interpretation of Summerfield's (1982) data suggests this is not the case.

Only empirical studies can resolve these issues. Primary among these would be a replication of Hirsh's (1959) study with a particular emphasis on the effect of varying instructions to the subject. Secondly, more standard psychophysical investigations are desirable in order to assess the discriminability of changes in tone-onset-time across the continuum, and how these changes are affected by the frequency and amplitude relationships between the two component tones of the complex. Initial forays in this direction have been made by Pastore et al. (1982).

These studies would be well worthwhile even if temporal order identification does not form the basis for the perception of voice onset time. As Hirsh originally pointed out, determining the order of auditory events is still an important aspect of auditory perception in general, with wide-ranging implications for understanding the perception of music and speech.

FOOTNOTES

1. Cutting and Rosner's (1974) study of the categorical perception of nonspeech "pluck"/"bow" and voiceless affricate/fricative continua, all of which were based on variations in rise time, was once considered influential in this regard. Unfortunately, their results have not withstood replication. See Rosen and Howell (in press) for a review.

2. Hirsh (1959), at least for the condition where the stimulus complex contained tones at 250 and 1200 Hz, used the phon scale to set the tones to be equal in loudness at a level of 80 phons. Assuming the tones in the other four two-tone conditions to be set at the same level, the intensity difference between the two tones would not have

been greater (and often smaller) than 2 dB (Fletcher and Munson 1933). In addition Hirsh varied the level of the 250 Hz tone, setting it to 60 and 70 phons, and found results "... not very different from the equal-loudness case." Note though, that in Pisoni's main experiment, the lower frequency tone was the more intense.

3. This may seem a little odd in that we are making the assumption that discrimination is predictable from identification or, in Macmillan, Kaplan, and Creelman's (1977) definition, that the continuum is categorically perceived! It is legitimate here because Pynn, Braida, and Durlach (1972) show that in experiments where the signals span a small range, discrimination distances inferred from identification experiments are close to those estimated directly. That Hirsh's (1959) stimuli do indeed only span a small range is attested to by the fact that the subjects never do better than about 95% correct at the extremes of the range, even though only two categories are involved.

4. Pisoni shows that the peak in the obtained discrimination function is well predicted from the labelling function using the so-called "Haskins formula". This formula can also be applied to Hirsh's labelling data, which then predicts a discrimination function peaked at 0 ms, just as our previous analysis did.

REFERENCES

1. Bock, R.D. and Jones, L.V. (1968). The Measurement and Prediction of Judgment and Choice. San Francisco: Holden-Day.
2. Cutting, J.E. and Rosner, B.S. (1974). Categories and boundaries in speech and music. Perception & Psychophysics 16, 564-570.
3. Fletcher, H. and Munson, W.A. (1933). Loudness, its definition, measurement, and calculation. Journal of the Acoustical Society of America, 5, 82-108. Also in: (1969) Forty Germinal Papers in Human Hearing, ed. J.D. Harris, pp. 175-184. Groton, Connecticut: The Journal of Auditory Research.
4. Hirsh, I. (1959). Auditory perception of temporal order. Journal of the Acoustical Society of America, 31, 759-767.
5. Howell, P. and Rosen, S. (1984) Natural auditory sensitivities as universal determiners of phonemic contrasts. In: Explanations for Language Universals, ed. B. Butterworth, B. Comrie, and O. Dahl. Berlin: Mouton. Also: Linguistics 21, 205-235.
6. Kaplan, H.L., Macmillan, N.A., and Creelman, C.D. (1978). Tables of d' for variable-standard discrimination paradigms. Behaviour Research Methods & Instrumentation, 10, 796-813.
7. Liberman, A.M., Cooper, F.S., Shankweiler, D.P., and Studdert-Kennedy, M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.
8. Liberman, A.M., Harris, K.S., Kinney, J.A. and Lane, H. (1961). The discrimination of relative onset time of the components of certain speech and nonspeech patterns. Journal of Experimental Psychology, 61, 379-388.
9. Macmillan, N.A., Kaplan, H.L., and Creelman, C.D. (1977). The psychophysics of categorical perception. Psychological Review, 84, 452-471.
10. Miller, G.A. (1956). The magical number seven plus or minus two, or, some limits on our capacity for processing information. Psychological Review, 63, 81-96.

11. Miller, J.D., Wier, C.C., Pastore, R.E., Kelly, W.J., and Dooling, R.J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. Journal of the Acoustical Society of America, 60, 410-417.
12. Parducci, A. (1965). Category judgment: A range-frequency model. Psychological Review, 72, 407-418.
13. Pastore, R.E., Ahroon, W.A., Baffuto, K.J., Friedman, C., Puleo, J.S., and Fink, E.A. (1977). Common-factor model of categorical perception. Journal of Experimental Psychology: Human Perception and Performance, 3, 686-696.
14. Pastore, R.E., Harris, L.B., and Kaplan, J.K. (1982). Temporal order identification: Some parameter dependencies. Journal of the Acoustical Society of America, 71, 430-436.
15. Pisoni, D.B. (1977). Identification and discrimination of the relative onset time of two-component tones: Implications for voicing perception in stops. Journal of the Acoustical Society of America, 61, 1352-1361.
16. Pisoni, D.B. (1980). Adaptation of the relative onset time of two-component tones. Perception & Psychophysics, 28, 337-346.
17. Pynn, C.T., Braida, L.D. and Durlach, N.I. (1972). Intensity perception: III. Resolution in small-range identification. Journal of the Acoustical Society of America, 51, 559-566.
18. Rosen, S. and Howell, P. (in press). Auditory, articulatory and learning factors in categorical perception. In: Categorical Perception, ed. S. Harnad. Cambridge University Press.
19. Stevens, K.N. (1981). Constraints imposed by the auditory system on the properties used to classify speech sounds: Data from phonology, acoustics and psycho-acoustics. In: The Cognitive Representation of Speech, ed. T.F. Myers, J. Laver and J. Anderson. Amsterdam: North Holland.
20. Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., and Cooper, F.S. (1970). The motor theory of speech perception: A reply to Lane's critical review. Psychological Review, 77, 234-249.
21. Summerfield, Q. (1982). Differences between spectral dependencies in auditory and phonetic temporal processing: Relevance to the perception of voicing in initial stops. Journal of the Acoustical Society of America, 72, 51-61.
22. Torgerson, W.S. (1985). Theory and methods of scaling. New York: John Wiley & Sons, Inc.

The Psychophysics of Speech Perception

edited by:

M.E.H. Schouten

Institute of Phonetics
University of Utrecht
Utrecht
The Netherlands

1987 **Martinus Nijhoff Publishers**
Dordrecht / Boston / Lancaster
Published in cooperation with NATO Scientific Affairs Division