A Range Effect in the Perception of Voicing*

Susan A. Brady† and Christopher J. Darwin‡†

ABSTRACT

The location of the voicing boundary in the perception of initial stop consonants is shown to vary according to the range of voice onset times used in a block of trials and according to the order in which blocks covering different ranges are presented. Although these range effects introduce methodological complications into the interpretation of adaptation experiments, they appear to be qualitatively different from adaptation effects and, it is suggested, may provide a metric for assessing the auditory tolerance of phonological categories.

INTRODUCTION

The numerical categories that subjects assign to particular stimuli along arbitrary dimensions, such as circle size or line length, are influenced by both the range and relative frequency of occurrence of stimulus values in the experiment (Helson, 1964). A particular stimulus will, for example, be assigned to a lower category when the range of stimuli used extends further to the high category end. The effects of frequency are more complicated but suggest a tendency for subjects to place equal numbers of stimuli in each category. Frequently occurring adjoining stimuli thus are allocated an undue wide range of categories (Parducci, 1974).

These effects of range and frequency are commonly found for dimensions that do not fall into "natural categories" (Rosch, 1973) and the question has been raised (Sawusch and Pisoni, 1974; Studdert-Kennedy, 1976) whether the perception of category boundaries in speech is immune from such effects. Changes in the relative frequency with which different stops, taken from either a voicing (Sawusch and Pisoni, 1974) or a place of articulation (Sawusch, Pisoni and Cutting, 1974) continuum, are played to subjects do not influence the position of their phoneme boundaries. On the other hand,


†University of St. Andrews, Fife, Scotland.

‡‡University of Sussex, Brighton, U.K.

Acknowledgment: Experiment I was supported by a grant from NICHD to Haskins Laboratories and Experiment II by a grant from the British SRC to Christopher J. Darwin.

similar frequency manipulations for fundamental frequency do move the point at which subjects change from using the label "high" to using "low" (Sawusch and Pisoni, 1974; Sawusch et al., 1974).

This suggested immunity is of considerable theoretical interest and potentially removes a number of methodological problems (Poulton, 1973) from the interpretation of, for example, adaptation experiments, whose results have generally been interpreted differently from range and frequency effects. However, it has been pointed out that range and frequency effects operate independently (Parducci, 1974), so it is possible that speech categories might be prone to range effects while being immune from frequency effects.

In the two experiments presented here, we first show that small but significant changes in the voicing boundary for stops can be obtained when the range of sounds presented within a block of trials is varied. We then go on to discuss the significance of this range effect in adaptation experiments and also to raise the possibility that range effects may be used to map the auditory tolerance of our internal phonological categories.

**EXPERIMENTAL DESIGN AND PROCEDURE**

In both experiments a synthetic voice onset time (VOT) continuum was used. An alveolar stop was synthesized on the Haskins Laboratories Parallel Formant Synthesizer before the diphthong /aI/ with 11 different VOT values from 5 to 55 msec in 5-msec steps, in order to give a continuum of sounds perceived as /daI/ with short VOT and /taI/ with long. The acoustic correlates of the change in VOT were a cut-back in first formant amplitude and a substitution of hiss for buzz excitation. In the first experiment each syllable was preceded by a carrier phrase "I say," but in the second experiment this was omitted. The two experiments were otherwise of identical design. Subjects listened to blocks of 40 trials in which the stimuli were drawn from five contiguous VOT steps: A(5-25 msec), B(15-35 msec), C(25-45 msec), D(35-55 msec). They also listened to an additional block covering the entire range of 11 steps: E(5-55 msec). Each block contained a random ordering of eight examples of each stimulus from the appropriate range. In each experiment, eight groups of four subjects listened to different orderings of these blocks; half the groups started with E and half finished with it, and within these halves each of the four groups took a different predecessor-balanced Latin-square order. The subjects, who for the first experiment were undergraduates at the University of Connecticut, and for the second were at the University of Sussex, were instructed to label each sound as being either more like a "d" or a "t." It was made clear to subjects that within a block of trials the proportion of either category could vary between zero and 100 percent. Subjects were encouraged to remember this and to judge each trial independently.

**Results**

Percent "d" responses for each VOT are shown in Figure 1 according to the range used in each block of trials, averaged across the four groups in each experiment who listened to the entire range block (E) prior to listening to the subset ranges (A, B, C, D) and also across those who listened to the E block last. There are significant differences between the responses given to
Figure 1: Percent "d" responses to stimuli differing in voice onset time between [daI] and [taI]. Each quadrant presents data from a different group of 16 subjects. Those in Experiment I heard the target syllable preceded by "I may ..."; those in Experiment II heard it in isolation. The two left-hand quadrants are for subjects who heard a block of trials covering the entire VOT range last, and the two right-hand ones for those who heard it first. The data points for the entire range are connected by a wide line. Subjects also heard blocks of trials in which the stimuli were from subranges. Their responses to these are shown by narrow solid (ranges A and C) or narrow broken (B and D) lines. Triangles mark data points at the ends of the subranges.
a particular VOT, depending on the range in which it occurred. These differences are, of course, larger at VOT values near the boundary. Statistical analysis is problematic because of necessarily missing data, but the following tests have been performed:

(1) Since ranges B, C and E all cover the three crucial middle VOT values of 25, 30, 35 msec, an analysis of variance is possible on this restricted set of data. This showed a significant effect of range on the total number of voiced responses for each of the four sets of data illustrated in Figure 1, each with $p < 0.005$. Figure 1 shows that this effect is due to a particular VOT stimulus receiving fewer voiced responses when it occurs as part of a block of trials in which the range covers shorter VOT values, than when the range covers longer VOT values. There is also an interaction of range effect with the order in which the blocks were taken ($p < 0.005$ for all except Experiment II with E last, which gave $p < 0.05$).

(2) Looking at the three middle VOT values separately, Friedman two-way analyses of variance showed significant effects of range on total voiced responses at better than the 1 percent level for the 25 and 30 msec VOTs in each of the two experiments when the entire range (E) was last, and also at better than the 2 percent level in Experiment II for each of the three middle VOTs when range E was last. No significant variation was found for these three individual VOTs in Experiment I when range E came first.

(3) Any change in the pattern of results between the two experiments, which differed in whether the precursor "I may ..." was present, was assessed by two analyses of variance similar to those in (1), but with Experiment I versus II as an additional dimension. For subjects receiving the entire range (E) last, there was only a weak interaction of experiment number with range on the number of voiced responses ($p < 0.05$), but for subjects who took the entire range first, experiment number gave a three-way interaction with range and the order in which blocks were taken ($p < 0.005$).

Averaging over block orders, the range effects thus appear to be rather larger and more reliable when the block with the entire range is presented last than when it comes first, and there are also significant influences on the range effect of the order in which A, B, C and D were presented. These interactions indicate that the extent of the range effect is not restricted to a particular block, but rather that the position of the category boundary is influenced by the range of preceding blocks as well as that of the present block. Since the present design confounds a block's predecessor with its serial position, these effects cannot be analyzed systematically here, but a working hypothesis suggested by the data is that the effective range is determined by all the preceding blocks, perhaps weighted in favor of the current block.

Interpretation of the three-way interaction of experiment number (or presence of precursor) with range and block order, like the interpretation of the two-way interaction between range and block order, is complicated by the confounding of a block's serial position with its predecessor; however,
inspection of the data suggests that it is mainly due to the sounds in range B. When there is a precursor, this range is heard as progressively more voiced the later in the experiment it is presented. When there is no precursor the same pattern occurs, except that when B is heard as the first block its sounds are consistently heard as more voiced than for later presentations.

There is thus some evidence that both a precursor and block order can influence the effect of stimulus range; but they are not sufficiently powerful to remove the effect, since we still find a significant (though reduced) range effect in the least favorable condition, when a precursor is present and the entire range is presented first.

Discussion

These two experiments give clear evidence that the perceived voicing of a sound depends quite markedly on the range of other sounds presented before it in an identification experiment. The more voiced the previous sounds are, the more voiceless they will appear. Similar effects of range are also present in the results of Lisker, Liberman, Erickson and Dechovitz [1975; Lisker, 1975, Figure 1]. There the VOT boundary in a /da/-/ta/ distinction was subject to the range of first formant transition durations used in the experiment. Again, their data indicated that the more voiced the companion stimuli, the more voiceless would a particular stimulus sound.

It might be argued that the results of our Experiment II could be interpreted as reflecting a general linguistic mechanism that compensates for the apparent rate of articulation of the speaker; this is unlikely because a constant rate precursor does not eliminate the effect (Experiment I) and because this hypothesis would not predict any range effect in Lisker's data. Linguistic categorizations, like other perceptual distinctions, rely on perceptual contrast and, it would seem, the decision mechanisms that register this contrast can be influenced by factors that are apparently linguistically irrelevant. Perhaps such flexibility should be welcomed in a communication system that must cope with the idiosyncrasies of individual speakers and with the varied distortions to which the speech signal is subjected.

The mean results shown in Figure 1 suggest that the range effects found here may be asymmetrical, with subjects being more willing to perceive as unvoiced a sound to the long-VOT end of a short-VOT range than to perceive as voiced a sound to the short-VOT end of the corresponding long-VOT range. An adequate statistical assessment of this is complicated by the interaction with block order and by the phoneme boundary not being exactly in the middle of the entire range, but if borne out in a more suitable experimental design, it would suggest that the extent to which range effects can influence phoneme boundaries is limited by the phonetic plausibility of the result. There is considerable variation of VOT in natural prestressed aspirated initial stops within the range of values that we have used for this experiment, and so range effects may be free to operate over these phonetically plausible values; but it is rare to find a natural apical unaspirated stop with a VOT of greater than about 20 msec before a stressed vowel, and perhaps this constrains the extent of range effects. This hypothesis of phonetic plausibility could perhaps also account for the much larger context effects noted by Eimas

101
(1962) in labeling data for triads of vowels than for triads of stops differing in place of articulation. The extent of range effects may thus provide a metric for assessing the auditory limits of our internal phonological categories.

The range effect found here and evident in other experiments might also contribute to the shift in stop voicing boundary found in adaptation experiments following repeated presentation of either a voiced or a voiceless stop, or some of the acoustic cues serving the voicing distinction [Eimas and Corbit (1973); Eimas, Cooper and Corbit (1973); Miller and Eimas (1975); Ades (1976)]. In particular, the early trials in the test phase of an adaptation experiment might form, with the adapting sound, a range that would differ depending on the adaptor. Although we cannot rule out the possibility of some of the adaptation effect being due to range, it is unlikely that all of the adaptation effect can be thus explained. The reason for this is that while adaptation effects are larger following adaptation to a voiceless stop [Eimas and Corbit (1973); Eimas et al. (1973); Miller and Eimas (1975)], our range effect seems to be greater for ranges extending into the voiced end of the voicing continuum. This difference in asymmetry may reflect a difference in underlying mechanisms, with adaptation being attributable to a change in the auditory representation of acoustic events, through adaptation of complex auditory feature detectors and range effects being attributable to a changed phonemic interpretation of a constant auditory representation. The greater phoneme boundary shift following voiceless adaptation could then be due to the voiceless adaptor reducing the sensitivity of a detector for long VOTs, which is perceptually more salient than the detector for low frequency first formant onset that might be reduced in sensitivity following adaptation to a voiced stop (Ades, 1976). The greater shift in phoneme boundary for a predominantly voiced than for a predominantly voiceless range, on the other hand, could be explained by the hypothesis of phonetic plausibility described earlier.

Thus, although we have shown that whatever natural categories we might possess for the voicing distinction are not immune from the biasing effects of range, nevertheless the size of the changes we find is quite small and they perhaps reflect in their magnitude the bounds of these internal categories.

REFERENCES


