Proceedings of The Institute of Acoustics

ADAPTATION OF PITCH CHANGE JUDGMENTS BY REPETITION.

W. A. AINSWORTH

UNIVERSITY OF KEELE

Introduction

It seems quite natural to invoke categorical judgments when discussing the perception of consonants because it is very difficult to produce a consonant with simultaneous closure of different parts of the vocal tract. The consonant /b/, for example, is produced with the lips whilst the consonant /d/ is produced with the tongue. To produce a consonant which is partly /b/ and partly /d/ is impossible. Therefore, if speech contains only unambiguous consonantal categories it would seem reasonable for the perceptual mechanism to reflect this, and in fact experiments with synthesised speech have confirmed categorical perception (Liberman et al., 1957).

With the perception of pitch, however, there are no such constraints to suggest categorical judgments. Fundamental frequency can be, and is, varied continuously by every speaker. It is only when speech recognition is considered that it is necessary to invoke categorical judgments.

Consider the word 'yes' spoken with a falling tone. It conveys the meaning that the speaker agrees with what has just been said. On the other hand, the word 'yes' spoken with a rising tone suggests that the speaker is questioning what has just been said.

At some stage in the processing of speech a decision has to be made about whether the fundamental frequency of an utterance is rising or falling. The question arises as to whether the mechanism underlying this decision is similar to the mechanisms involved in consonantal category judgments.

Method

The first experiment was designed to explore the ability of human listeners to judge whether the pitch of a word was rising or falling.

Nine versions of the word 'yes' were generated by means of a speech synthesis-by-rule system (Ainsworth, 1976). In each stimulus the fundamental frequency began at 120 Hz and ended at a value in the range 100 to 150 Hz. These stimuli were played 8 times in different randomised orders to a group of seven subjects who were asked to judge whether the pitch of each stimulus was rising or falling.

The results showed sharp perceptual boundaries in most cases. The boundary, however, was not located at the monotone position, but at some 10 Hz above this. Fourcin (1975) has obtained similar results with the diphthong 'oh', synthesised with changing fundamental frequency, being judged as a statement or a question.

ADAPTATION OF PITCH CHANGE JUDGMENTS BY REPETITION.

The next experiment was designed to discover whether this 'bias' of the perceptual boundary is a function of the rate of change of pitch or the extent of pitch change. To this end stimuli were synthesised having vowel durations of 160 and 640 msec. The experiment was repeated and almost identical values of the 'bias' were obtained (Table I), showing that it is the extent of the pitch change which affects the judgments. The fundamental frequency has to rise by 10 Hz or more for the pitch to be perceived as rising.

Initial frequency (Hz)	Duration (msec)	Bias (Hz)
120	160	10
120	640	9
259	160	19
259	640	15

Table I

Bias of the perceptual boundary from the monotone position as a function of fundamental frequency and vowel duration.

The experiment was repeated with the same vowel durations but with the fundamental frequency starting at 260 Hz and changing to some value in the range 220 to 330 Hz. Again a bias above the monotone position was observed for the perceptual boundary, and the values obtained for the two durations were not significantly different. The 'bias', however, was about twice as large as with the 120 Hz stimuli (Table I).

Adaptation Experiments

It is now well established that the perceptual boundaries between phonemes can be shifted by repetition of one of the stimuli. Lisker and Abramson (1967) showed that one of the most important features for distinguishing voiced plosives from voiceless ones is the magnitude of the delay between the plosive burst and the onset of voicing. This is known as the voice onset time (VOT). Eimas and Corbit (1973) showed that if a voiced plosive was repeated many times the perceptual boundary between voiced and voiceless plosives shifted along the VOT dimension in the direction away from the voiced plosive. Adaptation with voiceless plosives produced a shift in the opposite direction.

The next experiments were designed to discover whether the perceptual boundary between rising and falling pitch could be shifted by repetition of the stimuli in a similar manner. The word 'yes' was synthesised as before with the pitch changes shown in Table II. The perceptual boundary is shown with a slightly rising fundamental. The adapters were taken from the ends of the ranges: the 'falling' adapter had a fundamental which fell from 120 Hz to 100 Hz, 20 Hz below the monotone; and the 'rising' adapter had a fundamental which rose from 120 Hz to 150 Hz, 20 Hz above the unadapted perceptual boundary.

ADAPTATION OF PITCH CHANGE JUDGMENTS BY REPETITION.

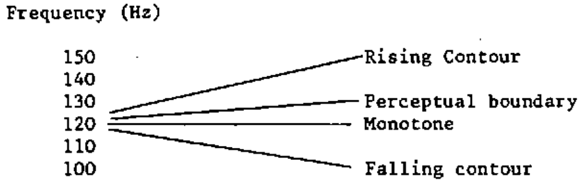

Table II

Frequency contours used in the adaptation experiments.

The positions of the unadapted boundaries for individual subjects were determined first. Then the experiments were repeated with the rising and falling adapters repeated 32 times before each test stimulus. In each session four runs with the same adapters were completed. These sessions took place at weekly intervals with the adapters chosen randomly. Eight runs were used to determine the position of each boundary.

These experiments were repeated with 'no' used as the carrier word. They were then repeated again with 'no' as the adapter and 'yes' as the test word, and vice versa. A summary of all the adaptation experiments is given by the first three columns of Table III.

Results

The relative positions of the adapted perceptual boundaries compared with the unadapted positions are shown in Table III for the various conditions.

Adapter	Contour	Test	Unadapted	Shift (Hz)	S.D.	Sign
Yes	Rising	Yes	Yes	5.7	5.8	5
No	Rising	No	No	5.9	5.0	6
Yes	Falling	Yes	Yes	14.0	6.5	6
No	Falling	No	No	17.6	5.3	7
No	Rising	Yes	Yes	1.7	7.1	0
Yes	Rising	No	No	4.3	5.3	4
No	Falling	Yes	Yes	8.1	5.3	5
Yes	Falling	No	No	8.9	5.6	5
---	---	Yes	No	3.0	5.3	2

Sign. = Number of subjects (maximum 7) for which the shift was significant at 0.1 level or better.

Table III

Table III Perceptual boundary shifts obtained with various frequency contours and carrier words.

ADAPTATION OF PITCH CHANGE JUDGMENTS BY REPETITION.

The differences in bias between the unadapted 'yes' and 'no' averaged 3 Hz, which was not significant at the 0.1 level (t - test) for five out of the seven subjects.

With 'yes' as the carrier word there was a shift of 5.7 Hz with the rising adapter and 14.0 Hz with the falling adapter. The falling adapter had a pitch change of only 20 Hz whereas the rising adapter had one of 30 Hz and this produced a smaller shift. Asymmetrical boundary shifts, however, are not uncommon in adaptation experiments (Cooper, 1974).

Similar results were obtained with 'no' employed as the carrier word (Table III). The phonemic composition and the acoustic structure of 'no' is very different from 'yes', so it would appear that the boundary shift effect is dependent only on the fundamental frequency contour. On the other hand, the results with the cross-series adapters showed that if the adapter and text words are different, the boundary shift is only about half as great as when they are the same. Again this is fairly typical of the results obtained from adaptation experiments with consonants (Eimas and Corbit, 1973).

Conclusions

The results obtained are best summarised by combining the 'yes' and 'no' data. The main conclusions are:

1) The bias is independent of the carrier word.
2) With the adapter rising or falling there is a significant shift, but it is about twice as great with the adapter falling.
3) In the cross-series conditions the shifts are about half as big as when the adapter and test words are the same.

The perceptual boundary shifts produced by adaptation are small, the experiments are boring for the subjects, and individual differences are large, but the differences between the various conditions are consistent. The results obtained suggest that the underlying mechanism responsible for pitch change judgments during speech perception has similar properties to the ones responsible for consonant categorisation.

References

W. A. AINSWORTH 1976 Proc. Inst. Acoustics, Autumn Meeting, 1976, 2-1-1 - 2-1-4. A versatile system for speech perception experiments.

W. E. COOPER 1974 J. Acoust. Soc. Am., 56, 617 - 627. Adaptation of phonetic feature analysers for place of articulation.

P. D. EIMAS and J. D. CORBIT 1973 Cognitive Psychol., 4, 99 - 109. Selective adaptation of linguistic feature detectors.

A. J. FOURCIN 1975 Speech perception in the absence of speech productive ability, in Language Deficits and Retardation, Butterworth.

A. M. LIBERMAN, K. S. HARRIS, H. S. HOFFMAN, and B. C. GRIFFITH 1957 J. Exptl. Psychol., 54, 358 - 368. The discrimination of speech sounds within and across phoneme boundaries.

L. LISKER and A. S. ABRAMSON 1967 Language and Speech, 10, 1 - 28. Some effects of voice onset time in English stops.