A model for dependencies in phonetic categorization

Roel Smits

Dept. Phonetics & Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, U.K.
Currently with Philips Research Labs, Building WY8-63, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

Abstract: A quantitative model of human categorization behavior is proposed, which can be applied to 4-alternative forced-choice categorization data involving two binary classifications. A number of processing dependencies between the two classifications are explicitly formulated, such as the dependence of the location, orientation, and steepness of the class boundary for one classification on the outcome of the other classification. The significance of various types of dependencies can be tested statistically. Analyses of a data set from the literature shows that interesting dependencies in human speech recognition can be uncovered using the model.

INTRODUCTION

Since the early days of research on speech perception there has been considerable interest in processing dependencies in human speech recognition (e.g. Miller and Nicely, 1955). The reason for this is that knowledge on such dependencies provides insight into fundamental aspects of human speech recognition, such as the decoding of coarticulated segments and the role of talker and speaking-rate information in phonetic perception.

The main experimental tool for carrying out such research has been the Garner paradigm (e.g. Eimas *et al*, 1979; Green *et al*, 1997). Although the Garner paradigm has proved a useful instrument for investigating processing dependencies in human speech recognition, it suffers from the problem that it is not based on an explicit quantitative model of the psychological mechanisms involved. Distinct types of perceptual and cognitive processing dependencies (Ashby and Townsend, 1986) are confounded in the Garner paradigm. As a result, the experimenter can only draw conclusions concerning the significance and directionality of dependencies in processing the two dimensions, and does not achieve a quantitative description of the processing involved.

In this paper a new quantitative model for dependencies in phonetic categorization is defined which does not suffer from these drawbacks. The model describes categorization dependencies in a 4-alternative forced-choice experiment involving two binary distinctions d_1 and d_2 (for example, voicing and place with responses /b, d, p, t/), and applies to categorization data (counts) rather than reaction times.

DEFINITIONS

First of all, we define dependence and independence in a probabilistic sense. It is assumed that the subject classifies the stimuli along two dimensions d_1 and d_2, each of which can have 2 values (+, -). The subjects' categorization along d_1 and d_2 is defined as independent if the joint probability of categorization along d_1 and d_2 equals the product of the individual probabilities for categorization along d_1 and d_2:

Independence: $$p(d_1+, d_2+) = p(d_1+)p(d_2+) \quad (1)$$

and analogously for negative values of d_1 and d_2. For example, if stops are classified independently along dimensions voice and place: $p(/\text{b}/) = p(\text{voiced}, \text{labial}) = p(\text{voiced})p(\text{labial})$. In a similar fashion we define:

Dependency 1: d_2 depends on d_1: $$p(d_1+, d_2+) = p(d_1+)p(d_2+ \mid d_1+) \quad (2)$$

Dependency 2: d_1 depends on d_2: $$p(d_1+, d_2+) = p(d_2+)p(d_1+ \mid d_2+) \quad (3)$$

We now face the problem that (2) and (3) always hold and are indistinguishable on the basis of the response pattern to a single stimulus. However, with extra assumptions on the underlying perceptual space and categorization functions, we can distinguish between the two dependency directions for a *set* of stimuli. It is assumed that x and y are axes in perceptual space associated with d_1 and d_2 and that the classification functions along these axes are logistic. In the case of the d_2 categorization depending on the d_1 categorization, this leads to the following model.

$$\log \frac{p(d_1+)}{p(d_1-)} = x \tag{4}$$

$$\log \frac{p(d_2+|d_1+)}{p(d_2-|d_1+)} = y + (c_0 + c_x x + c_y y) \tag{5}$$

$$\log \frac{p(d_2+|d_1-)}{p(d_2-|d_1-)} = y - (c_0 + c_x x + c_y y) \tag{6}$$

where the parameters c_0, c_x and c_y are associated with a dependency of the *position, orientation* and *steepness* of the d_2 boundary on the outcome of the d_1 categorization.

The model has been implemented for categorization data obtained with 2-D stimulus continua, using the additional assumption that psychological axes x and y are linearly related to the physical parameters used in the construction of the stimulus continuum.

MODEL FIT ON DATA FROM LITERATURE

Whalen (1989) created a 2-dimensional stimulus continuum by varying the F1 steady-state frequency and the steady-state duration of a synthetic CVC word in 3 steps and 10 steps, respectively. Each of the 30 resulting stimuli was presented 20 times to each of 16 subjects. Subjects were instructed to categorize the stimuli as *bad, bat, bed,* or *bet*. The dependency model defined above was fitted on Whalen's data, and the significance of various dependency parameters was tested statistically using the G^2 statistic. For the resulting best model (G^2=144) the categorization of the final consonant was dependent on the categorization of the preceding vowel. Parameter values were c_0=-0.18, c_y=0.11, while c_x was not significantly different from zero. Conceptually, these parameter values mean: `if ba then steep t-d boundary at high duration and low F1;`
`if be then shallow t-d boundary at low duration and high F1.`
These mechanisms can be verified using Figure 1, which gives a graphical representation of the model shape.

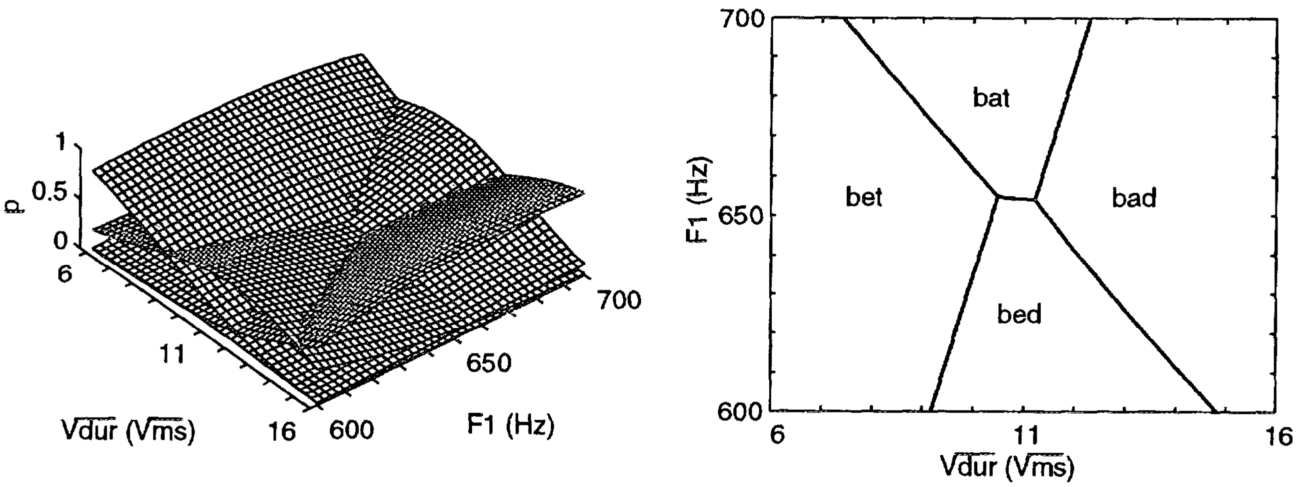

FIGURE 1. Shape of model fitted on data of Whalen (1989). Left panel: probability surfaces, right panel: territorial plot.

REFERENCES

Ashby, F.G., and Townsend, J.T. *Psychological Review* **93**, 154-179 (1986).
Eimas, P.D., Tartter, V.C., Miller, J.L., and Keuthen, N.J. *Perception & Psychophysics* **23**, 12-20 (1979).
Miller, G.A., and Nicely, P.E. *J. Acoust. Soc. Am.* **27**, 338-352 (1955).
Whalen, D.H. *Perception & Psychophysics* **46**, 284-292 (1989).