Discrimination functions predicted from categories in speech and music

JAMES E. CUTTING
Wesleyan University, Middletown, Connecticut 06457
and
Haskins Laboratories, New Haven, Connecticut 06511

and

BURTON S. ROSNER
University of Pennsylvania
Philadelphia, Pennsylvania 19174

Sawtooth waves differing only in rise time are identifiable as plucked or bowed notes from a stringed instrument. We previously reported (Cutting & Rosner, 1974) that these nonlinguistic sounds are perceived categorically. We also synthesized a continuum of speech sounds by varying only rise time. Listeners identified these sounds as /tʃa/ or /ʃa/, as in CHOP and SHOP, respectively, and perceived them categorically as well.

Our criteria for categorical perception were those suggested by Studdert-Kennedy, Liberman, Harris, and Cooper (1970): (a) "peaks" of high discriminability between stimuli in restricted regions along the dimension studied, (b) "troughs" of discrimination performance near chance in regions on either side of the peak, and (c) correspondence between peaks and troughs and the course of identification functions, with peaks occurring at identification boundaries and troughs occurring within each perceptual category. Categorical perception is therefore revealed by a particular combination of results from identification and discrimination tasks. This convergence between identification and discrimination is unusual: A listener generally can discriminate many more stimuli than he or she can identify absolutely (see, for example, Miller, 1956).

The correspondence between identification and discrimination can be tested quantitatively. Discrimination performance can be predicted from identification data by assuming that discrimination is no better than identification. To the extent that obtained and predicted discrimination scores do not differ significantly, categorical perception has occurred.

Our previous paper described such agreement between obtained and predicted discrimination scores for both the linguistic and musical sounds (Cutting & Rosner, 1974). Unfortunately, the predicted functions were derived through an incorrect formula. This note corrects that error.

To predict discrimination from identification of a two-category continuum, the correct formula for an ABX discrimination task is:

\[ P(c) = \frac{1}{2} (1 + p_1 - p_2)^2 \],

(1)

where \( P(c) \) is the probability of a correct discrimination, \( p_1 \) is the probability of assigning Stimulus A to one of the categories, and \( p_2 \) is the probability of assigning Stimulus B to that same category. The original formula for the three-category case published by investigators at the Haskins Laboratories (Liberman, Harris, Hoffman, & Griffith, 1957) is incorrect; Pollack and Pisoni (1971) give proper formulae for both two- and three-category continua. We will refer to Equation 1 as the Haskins prediction.

Typically, obtained discrimination functions, even for stop consonants, systematically exceed predicted functions by as many as 10 percentage points at each comparison along the stimulus array. Thus, the strongest possible relationship between identification and discrimination is not realized (see also Barclay, 1972; Pisoni & Lazarus, 1974; and Pisoni & Tash, 1974). The discrepancy between obtained and predicted discrimination functions is even larger for more "continuously" perceived stimuli such as vowels (Pisoni, 1973, 1975). By further developing a model that Fujisaki and Kawashima (1970) formulated, Pisoni added a correction factor to prediction formulae such as (1). This factor is based on the asymptotic through discrimination value, it raises the predicted functions by several percentage points, and it can be interpreted as measuring short-term auditory storage for differences between two stimuli identified alike. For a two-category continuum in an ABX task, the proper Fujisaki-Kawashima prediction formula is

\[ P(c) = \frac{1}{2} [(p_1 - p_2)^2 + p_1(1 - p_2) + p_2(1 - p_1)] + T \]

(2)

where \( P(c) \), \( p_1 \), and \( p_2 \) are the same as in Equation 1 and \( T \) is the asymptotic trough value of the obtained discrimination function. If \( T = 0.50 \), Equation 2 reduces to Equation 1. Like the Haskins prediction formula, the Fujisaki-Kawashima formula has suffered the misfortune of appearing incorrectly in print (Pisoni, 1971, p. 44; Pisoni, 1975, p. 13).1

Using the correct formulae, we have recomputed both the Haskins and the Fujisaki-Kawashima predictions for our data on discrimination of sawtooth waves and of affricate-fricative speech syllables. Pre-

- 1We thank Neil Macmillan for pointing out our error to us and for suggesting how it occurred.

87
Table 1
Obtained and Correctly Predicted Discrimination Values for Stimuli Differing in Rise Time

<table>
<thead>
<tr>
<th></th>
<th>0-20</th>
<th>10-30</th>
<th>20-40</th>
<th>30-50</th>
<th>40-60</th>
<th>50-70</th>
<th>60-80</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sawtooth Wave Stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obtained</td>
<td>61</td>
<td>64</td>
<td>72</td>
<td>78</td>
<td>58</td>
<td>60</td>
<td>59</td>
</tr>
<tr>
<td>Haskins Predicted</td>
<td>50</td>
<td>51</td>
<td>60</td>
<td>67</td>
<td>56</td>
<td>52</td>
<td>51</td>
</tr>
<tr>
<td>Fujisaki-Kawashima Predicted</td>
<td>58</td>
<td>59</td>
<td>65</td>
<td>71</td>
<td>61</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td>Speech Stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obtained</td>
<td>61</td>
<td>58</td>
<td>59</td>
<td>70</td>
<td>76</td>
<td>61</td>
<td>58</td>
</tr>
<tr>
<td>Haskins Predicted</td>
<td>51</td>
<td>53</td>
<td>55</td>
<td>62</td>
<td>64</td>
<td>51</td>
<td>50</td>
</tr>
<tr>
<td>Fujisaki-Kawashima Predicted</td>
<td>58</td>
<td>59</td>
<td>60</td>
<td>67</td>
<td>68</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td><strong>Experiment II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sawtooth Wave Stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obtained</td>
<td>61</td>
<td>55</td>
<td>66</td>
<td>72</td>
<td>47</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td>Haskins Predicted</td>
<td>50</td>
<td>50</td>
<td>66</td>
<td>73</td>
<td>54</td>
<td>51</td>
<td>50</td>
</tr>
<tr>
<td>Sine Wave Stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obtained</td>
<td>54</td>
<td>49</td>
<td>56</td>
<td>68</td>
<td>56</td>
<td>58</td>
<td>53</td>
</tr>
<tr>
<td>Haskins Predicted</td>
<td>50</td>
<td>51</td>
<td>63</td>
<td>68</td>
<td>54</td>
<td>54</td>
<td>53</td>
</tr>
</tbody>
</table>

Note—The original predicted functions that appear in Cutting and Rosner (1974) are incorrect.

dictions were made for each individual listener, then averaged functions were obtained from the individual functions, as Pisoni (1971) suggests. Table 1 shows averaged obtained and predicted discrimination scores.

The predicted functions in Table 1 are farther below the obtained functions than were those originally published (see Tables 1 and 2 in Cutting & Rosner, 1974). Nevertheless, the discrepancies between predicted and obtained scores here are not marked. Goodness-of-fit measures calculated from individual obtained and Haskins-predicted scores revealed no significant differences (see Pisoni, 1971, p. 20), although the observations per comparison may be too few to make small differences statistically reliable. The fit between the data and the correct predictions still supports our prior conclusion that musical stimuli and affricate-fricative consonants differing in rise time are each perceived categorically. Subsequent experiments have provided confirmation: Cutting, Rosner, and Foard (1976) have demonstrated that the musical sounds are perceived as categorically as stop consonants in Pisoni’s (1971, 1973) variable-interval AX discrimination task.

In summary, this note presents correct predicted discrimination functions for data previously published (Cutting & Rosner, 1974). The corrections leave the principal conclusion of that study unchanged: Nonlinguistic and linguistic stimuli synthesized with different rise times are perceived categorically. In addition, this note provides correct formulae for predicting discrimination functions. Several previous sources for the formulae are in error.

REFERENCES


MILLER, G. A. The magical number seven, plus or minus two, or some limits on our capacity for processing information. Psychological Review. 1956, 63, 81-96.


NOTES

1. Page numbers for Pisoni (1971) refer to a version published as a supplement to the Haskins Laboratories Status Report on Speech Research.

2. The trough value T was not stable for individual listeners; we assumed it to be 0.60 for all listeners for both sets of stimuli represented in Table 1.

(Received for publication June 21, 1976; accepted June 23, 1976.)