Categorical Perception Along an Oral-Nasal Continuum*

Roland Mandler+

ABSTRACT

Dental and labial nasal consonants were constructed using two methods of synthesis, one employing the nasal branch resonances, and one the oral branch resonances of the OVE III in simulation of period of closure. Oral-nasal continua were generated for both places (/da/ to /na/ and /ba/ to /ma/) for both methods of synthesis. Identification and same-different discrimination tests from all four resulting sets were administered to thirteen subjects. Their responses yielded strong evidence for categorical perception along the oral-nasal dimension.

Extensive analysis of the structure of nasal consonants by Fujimura (1962) and other researchers has revealed the predominant cue value of two basic components: one, the presence of a low amplitude noise through the period of closure, and two, a stoplike transition following the closure. Using a tape-splicing technique, Malécot (1956) confirmed that place cues were carried in the transition and nasality was cued by the low amplitude noise through the period of closure. A study by Liberman, Delattre, Cooper, and Gerstman (1954) using synthetic speech employed identical transitions for nasals and oral stops to get labeling judgments for continua across place of articulation. Both sets yielded good identification. These and other perception experiments suggested that listeners perceived such stimuli categorically, that is, distinguished stimuli across phonetic categories but not within categories, despite identical acoustic variations. The present experiment was undertaken to determine whether categorical perception could be evidenced along an oral-nasal continuum.

METHOD

Stimulus Specifications

The specifications of the OVE III serial synthesizer at the Haskins Laboratories allowed for two methods of construction of nasals from stops:

*Presented at the 91st meeting of the Acoustical Society of America, 4 April 1976, Washington, D.C.

+Also University of Connecticut, Storrs.

the first method, henceforth referred to as the oral branch method, simulates nasals on the oral branch of the synthesizer by making use of wider bandwidth settings for the first and second formants through the period of closure; the second method, henceforth referred to as the nasal branch method, preceded and overlayed oral branch stop transitions with the output of the nasal branch of the synthesizer, with one variable formant and a number of higher fixed formants.

Consonant-vowel nonsense syllables in configuration were chosen as the basic stimuli of the experiment. The neutral vowel /a/ was employed for both methods, with continua constructed for bilabial (that is, /ma/ to /ba/) and alveolar (/na/ to /da/) places of articulation. Figure 1 illustrates the extreme ends of the bilabial continua for both methods, with shading through the portions varied.

![Diagram of oral and nasal branch methods](image)

**Figure 1:** Schematic diagram of extreme nasal bilabial stimuli using both methods. Shading indicates portions varied through continua.

The extreme nasal stimulus of the oral branch method bilabials had the following structure: an 80-msec period of closure with $F_1$ at 240 Hz, $F_2$ at 1000 Hz, and $F_3$ at 2600 Hz, followed by a 40-msec transition to the vowel steady-state values of 820 Hz for $F_1$, 1180 Hz for $F_2$, and 2630 Hz for $F_3$. The vowel duration was 280 msec, resulting in a total stimulus length of 400 msec. Over the nine stimuli of the continuum, three parameters were varied in equal
steps during the period of closure: oral branch amplitude was varied from 0 through 24 dB relative to the amplitude of the vowel steady-state; F1 bandwidth was varied through the range 205 to 70 Hz; F2 bandwidth was varied through the range 350 to 80 Hz. The alveolar set consisted of the same acoustic variations with differences only in the initial formant frequencies, namely F2 at 2000 Hz and F3 at 2800 Hz.

The extreme nasal stimulus of the nasal branch method bilabials had the following structure: nasal branch excitation with its lowest resonance at 240-Hz through the 80-msec period of closure and the 40-msec stop transition; oral branch resonances initiated the transition from values of 240 Hz for F1, 1000 Hz for F2, and 2200 Hz for F3. The schematic does not illustrate the fixed upper formants of the nasal branch to prevent confusion with the upper formants of the oral branch. Only nasal branch amplitude was varied in this set, through the range 0 dB to -14 dB relative to the amplitude of the vowel steady-state. This variation resulted in an eight-member continuum. The alveolar set contained the same nasal formants and variations, with initial oral branch formant values identical to those of the oral branch method alveolar set.

**Procedure**

Pilot free-choice labeling tests were given to 20 subjects to determine that end points represented the nasals and stops intended, and to assure that only those two categories were perceived through all continua.

Final forced-choice identification tests, one randomly arranged test for each continuum, consisted of four presentations of each stimulus. Each presentation contained two samples of the stimulus with a one-second interstimulus interval. The interval between presentations was four seconds. Thus, each test set for the oral branch method contained 36 presentations, while each test set for the nasal branch method contained 32 presentations.

Same-different discrimination tests, one for each of the four continua, had the following form: as sames, four randomly arranged pairs of each stimulus were included; as differents, four randomly arranged pairings of adjacent stimuli, two in the order AB and two in the order BA, appeared. The interstimulus interval was 500 msec, and the interval between pairs was four seconds. The resulting oral branch method test sets consisted of 64-pair presentations for each place, while the nasal branch method sets consisted of 56 pairs each.

**Subjects**

Fifteen paid subjects, all University of Connecticut students, took all eight tests. All were native speakers of American English who claimed normal hearing ability and phonetic naiveté. Eight were right-handed males and seven were right-handed females. The test sets were presented binaurally through headphones in a soundproofed room. Up to five subjects took the tests at one time. All four identification tests were presented first; the four discrimination tests followed.
RESULTS

Figure 2 shows bilabial identification and discrimination results on the oral branch method for subject DC. This subject was given the tests again.

Figure 2: Oral branch method bilabial results for subject DC.
later in order to get more reliable identification and discrimination curves. Each point on the labeling graphs thus represents eight judgments, while each discrimination point represents sixteen responses. Standard methods of data analysis were used.

The labeling crossover occurs in the region of stimulus 6. The discrimination curve rises from a baseline value of approximately 50 percent to a peak of 75 percent in the region between stimuli 6 and 7. Thus, stimuli 1 through 5 formed the nasal category, and stimuli 7 through 9 made up the stop category for this subject. The alveolar data given by DC showed a similar pattern of agreement between labeling crossover and discrimination peak. Crossover was at stimulus 6 and the discrimination peak of 82 percent occurred for the pairing of stimuli 6 and 7.

Figure 3 shows the nasal branch method bilabial results for the same subject. The labeling crossover corresponds to stimulus 5 (which had nasal amplitude of -8 dB relative to the vowel steady-state amplitude). The discrimination peak of 89 percent corresponded to the pairing of stimuli 5 and 6. Again, good agreement is evident between crossover and discrimination peak. The alveolar data for this stimulus type showed a labeling crossover between stimuli 4 and 5, with a discrimination peak of 89 percent in that same region.

Figure 4 depicts discrimination responses for all four test sets given by DC. Correspondences across place for peak and baseline data show good consistency. Notable in the oral branch results is the fact that the slope toward the peak from the nasal category is greater than that away from the peak toward the oral-stop category. This suggests that a peculiarity of the method introduced greater cue value for within-category discrimination of oral stops, than for nasals.

The nasal branch method did not yield consistent correspondences between peak values across place. The peak for the bilabial set peak occurred in the region of stimuli 5 and 6, while the alveolar peak was in the region of stimuli 4 and 5, despite the fact that the acoustic variations were identical. Both sets, nonetheless, yielded well established nasal and stop categories at either end.

Group data for both stimulus types shows essentially the same patterns of distribution as were evident in this individual's responses. Perceptual variation across subjects causes a wider region of indecision around crossovers and wider and somewhat depressed discrimination peaks. What is apparent in even the group data, however, is the consistent correspondence between crossovers and discrimination peaks. These correspondences offer substantial evidence for categorical perception along an oral-nasal continuum.

A relevant problem brought out by the results of this experiment concerns a hypothesis of Fujisaki and Kawashima (1968, 1969, 1970). They proposed that consonants are perceived categorically, and vowels less so, due to the acoustically transient character of the consonantal cues. The continua of this study, however, relied on an 80-msec steady-state noise with varying amplitude for cue value. The categorical perception observed here therefore cannot be attributed to transience of the distinctive oral-nasal cue.
Figure 3: Nasal branch method bilabial results for subject DC.
Figure 4: Discrimination results for all four test sets given by DC.
REFERENCES


