1. Purpose

Stop and nasal consonants are both produced by occluding the vocal tract in the oral cavity and then releasing the constriction. The occlusion during a nasal consonant coincides with a lowering of the velum, creating a coupling between the main vocal tract and the nasal passages, and allowing for air flow and sound to exit at the nares. Although coupling is required for a nasal consonant to be produced, the minimum nasal cavity and then releasing the constriction. The occlusion during a nasal consonant modifies the area function.

The specific aim of the study was to determine the nasal coupling area at which the perception of a stop consonant switches to its nasal cognate.

2. Method

2.1 Speech production model

A model of the speech production system (Figure 1) was used to generate a collection of synthetic VCV utterances based on time-varying changes of the vocal tract area function (Stoy, 2013; Story & Bunton, 2019).

Figure 1: Area function model of the vocal tract with coupling to nasal tract. Nasal coupling also modifies the area function.

- The V1 for all stimuli was an /a/ vowel, and V2 was one of the four corner vowels /i, a, n, u/. The time course of each vowel is shown in Figure 2a as the gray "event" functions.
- For each V1CV2, the consonant consisted of an occlusion of the vocal tract in the labial, alveolar, or velar region accompanied by a time-varying change in nasal coupling area. The consonant time course is shown in Figure 2a as the pink event function.
- Plotted in the upper part of Figure 2a are eight time-varying functions that dictated the magnitude of the area coupling the main vocal tract to the nasal passages during production of a given VCV, where the maximum value of the coupling ranged from 0.1 cm². In total, 96 V1CV2 utterances were generated (4 V2 × 3 consonant types × 8 coupling area time functions). An example time-varying area function for /al/ or /als/ is shown in Figure 2b.

Figure 2: (a) Time course of vowel and consonant perturbations of the vocal tract shape (lower panel) for production of synthetic VCVs, along with accompanying time-varying change in nasal coupling area (upper panel). (b) Example time-varying area function generated by the event function for /als/ or /als/ (depending on nasal coupling area).

2.2 Demonstration of V1CV2 continua

- Figure 3 shows time-varying vocal tract configurations for the bilabial, alveolar, and velar series where the initial and final vowels are both /a/. The gray triangle in the velar region indicates the location of the nasal coupling port.
- The waveforms and spectrograms in each row represents the endpoints of the nasal coupling continuum: max[A2(t)] = 0 cm² in the middle column, and max[A2(t)] = 0.1 cm² in the right column. The red lines in the spectrograms indicate the time variation of the nasal coupling.

Time-varying vocal tract

max[A2(t)] = 0 cm²
max[A2(t)] = 0.1 cm²

Figure 3: Time-varying vocal tract configurations for bilabial, alveolar, and velar series of /a/ and /als/, and corresponding waveforms/spectrograms for the non-nasal condition, and the maximum nasal condition.

2.3 Listening experiment

- Responses in a perceptual experiment were collected from 14 listeners recruited from the student population at the University of Arizona. Each listener was a native speaker of American English and passed a hearing screening of 20 dB HL (American Speech-Language-Hearing Association, 1997). All procedures were approved by the Institutional Review Board at the University of Arizona.
- Individual listeners were seated in a sound booth where audio samples were presented over a loudspeaker placed one meter in front of the listener (Yamaha HSS studio monitor). The experiment was controlled by using the ALVIN interface (Hil-lerbrand and Guyvert, 2005), and each stimulus presentation consisted of one of the synthesized V1CV2.
- Listeners were instructed to identify the consonant using a forced-choice paradigm, where the computer screen displayed buttons for the six consonants /b, d, g, m, n, /ŋ/. Each stimulus was randomly presented four times, and listeners were allowed to replay any particular stimulus if needed.

Figure 4: Identification functions pooled from listener responses.

3. Results

Identification functions based on the pooled responses from all listeners are shown in Figure 4.

- In each panel, the black and red curves represent the identification of the non-nasal and nasal cognates, respectively, and are plotted as functions of the nasal coupling area achieved during consonant portion of the V1CV2. Any responses that were not the target consonants for a particular stimulus are indicated by the curves with open circles and the consonant identified is shown as a letter symbol.
- The identification functions in the first column were obtained by combining responses across all four V2 vowel contexts. The remaining four columns are based on the separate /i, a, n, u/ V2 contexts.
- Using a 50 percent criterion, the thin vertical line in each plot indicates the nasal coupling area at which the listener responses switched from identification of a stop to identification of a nasal.

<table>
<thead>
<tr>
<th>All vowel contexts</th>
<th>/iC1</th>
<th>/iC2</th>
<th>/iC3</th>
<th>/iC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Discussion

- Across all stimuli, the minimum nasal coupling value at which the response switched from stop to nasal was 0.036 cm² and the maximum was 0.057 cm².
- Typically, the smallest nasal coupling values at which a stop switched to a nasal was found for the stimuli with an alveolar consonant (second row).
- The vowel context in which the stop to nasal switch required the largest nasal coupling values was /iC4/. Perhaps this is because an /i/ vowel typically requires a larger amount of nasal coupling to be perceived as being nasalized than other vowels (cf., Bunton & Story, 2012).
- The listener responses for the velar consonants (third row) indicated a high degree of confusion with other consonants, likely because embedding /i/ in a V1CV2 produces an unusual combination of phonetic segments. Nonetheless, the data still indicate clear crossover points at nasal coupling values aligned with the other consonants.

Acknowledgements

Research supported by NIH R01-DC011275 and NSF BCS-1145011

Brad H. Story & Kate Bunton
Speech, Language, and Hearing Sciences
The University of Arizona