EFFECTS OF FACIAL PARALYSIS AND PRESENTATION MODE ON PERCEPTUAL-ACOUSTIC MEASURES OF CONSONANT PLACE

Monica A. Rohlfs & Megan M. Hodge
Department of Speech Pathology and Audiology
University of Alberta, Edmonton, Alberta T6G 2G4

1. INTRODUCTION

Listeners use both acoustic and visual cues to identify speech. Discrepancies between these cues increase the frequency of listener misidentifications of the speech signal[1]. A speaker with bilateral facial paralysis (BFP) may present conflicting visual and acoustic cues when producing bilabials. That is, despite the absence of lip movement because of facial paralysis, the speaker may compensate with the tongue and jaw to produce acoustic signals that listeners identify as bilabial sounds[2]. As the second formant frequency (F2) has been established as a primary cue to place of articulation[3], it was hypothesized that when sounds produced by a speaker with BFP are identified by listeners as bilabials, the acoustic cues to articulatory place contained in these sound productions are consistent with those normally associated with labial place of articulation.

The objectives of this study were:

1. To determine if listener identification of bilabial and alveolar consonants produced by a child with BFP and a child with normal facial muscle function (NFMF) was affected by mode of stimulus presentation, i.e., an auditory-only (A) versus an auditory-visual (A-V) signal. For the child with NFMF it was hypothesized that consonant identification scores would be high in both A and A-V conditions, with scores in the A-V condition expected to be slightly higher as there are more cues available for signal decoding. Conversely, consonants produced by the child with BFP were expected to be identified correctly with greater frequency in the A condition, due to conflicting auditory and visible articulatory cues in the A-V condition.
2. To determine if F2 coordinates and corresponding locus equations[4] of correctly identified bilabial and alveolar consonants in the A condition for the child with BFP were consistent with those produced by the child with NFMF. It was hypothesized that misidentifications would occur when F2 values were not consistent with those normally expected for the target place.

2. METHOD

Two girls, one with BFP and one with NFMF, were videotaped as they read CV words embedded in a carrier phrase. The consonant /b/ or /d/ was followed by one of the vowels /i, I, e, æ, ʌ, ɛ, a, o, u/, as in Sussman et al.[4]. Ninety utterances (2 consonants X 9 vowels X 5 repetitions) were recorded per child. Digital movie files of each utterance were created using a Macintosh Power PC 8500 and Avid VideoShop 3.0 software. HyperCard 2.1 was used to create a program that randomly presented the movie files to listeners and recorded and analyzed listeners' responses. After each item was presented, listeners selected the consonant that they perceived (either /b/ or /d/) by "clicking" on the appropriate letter choice displayed on the computer monitor. Listening sessions were conducted in a sound booth. For the A condition, only the audio track of the movie files was presented. Listeners were 47 speech-language pathology graduate students who were randomly assigned to the A (n=25) or A-V (n=22) condition. Within each condition, presentation order of the speaker (BFP and NFMF) was counterbalanced.

3. RESULTS

A three-way ANOVA tested the factors facial muscle function (BFP and NFMF), place of articulation (/b/ and /d/) and presentation mode (A and A-V). Each factor had a significant main effect and a significant three-way interaction was obtained ($p<.0001$). The child with NFMF had a higher number of correctly identified phonemes (b=99%; d=100%) than the child with BFP, regardless of consonant place and presentation mode. The child with BFP had a higher number of accurate identifications for /d/ (95%) than /b/ (30%), and a significantly greater number of correct identifications for /b/ in the A (51%) than the A-V (10%) mode. Acoustic analyses of the stimulus words were completed on a 486 personal computer, using CSpeech 4.0. F2 onset and offset were calculated for each "correctly identified" (validated) CV token, i.e., those CVs where at least 75% of listeners correctly identified the target consonant. Regression lines (locus equations) for each consonant place were generated following procedures outlined by Sussman et al.[4]. For both children's productions, /b/ regression lines had steeper slopes, while /d/ regression lines had higher y-intercepts. For the same consonant place, regression line slope was steeper, and y-intercept was higher for the child with BFP. Non-validated stimulus items were present only for the child with BFP. Her locus equation for the non-validated /b/ tokens had a slightly higher y-intercept but similar slope compared to the validated tokens.

Acoustic analysis revealed that the child with BFP produced some consonants in the CV syllables that were perceived by listeners as /b/ in the A-only condition and that had F2 values consistent with those expected for the bilabial place of articulation. These CVs tended to contain back vowels. However, in the A-V condition, her lack of lip movement caused listeners to perceive the CV tokens as starting with the alveolar consonant /d/, even when the F2 characteristics in the acoustic signal resembled those for the bilabial /b/. Thus, the visual cue for place "overrode" the conflicting acoustic cue, resulting in misperception of the intended consonant. The stimulus presentation procedures developed for this project show promise for measuring differential effects of auditory versus auditory-visual cues on speech identification scores of disordered speakers, and in the case of speakers with BFP, the effects of surgical or other treatments on speech intelligibility scores.

4. REFERENCES

1. McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.
2. De Feo, A. B. & Schaeffer, C. M. (1983). Bilateral Facial Paralysis in a preschool child: Oral-facial and articulatory characteristics (A case study). In W. R. Berry (Ed.), *Clinical Dysarthria.* College Hill Press, San Diego.
3. Borden, G. J., Harris, K. S. & Raphael, L. J. (1994). *Speech science primer: Physiology, acoustics, and perception of speech.* Baltimore: Williams & Wilkins.
4. Sussman, H. M., Fruchter, D. & Cable, A. (1995). Locus equations derived from compensatory articulation. *JASA, 97,* 3112- 3124.