Speech Compensation in Persons who Stutter: Acoustic and Perceptual Data

Aravind K. Namasivayam[1,2] and Pascal van Lieshout[2,3,4,5,6]

[1]The Speech and Stuttering Institute, Toronto, ON, Canada, M3B 3M4, e-mail:a.namasivayam@utoronto.ca
[2]Oral Dynamics Lab, Dept. of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada, M5G1V7
[3]Dept. of Psychology, Human Communications Lab, University of Toronto. 3359 Mississauga Road, L5L1C6
[4]Institute for Biomaterials & Biomedical Engineering, 164 College Street, University of Toronto, ON, Canada M5S 3G9
[5]Toronto Rehabilitation Institute, 550 University Ave. Toronto, Ontario M5G 2A2
[6]Graduate Department of Rehabilitation Science, 500 University Ave. Toronto, Ontario M5G 1V7

1. INTRODUCTION

According to the Speech Motor Skills (SMS) approach stuttering dysfluencies are viewed as audible manifestations of errors in motor control arising from limitations in speech motor skill (Van Lieshout, Hulstijn, & Peters, 2004). Within this framework persons who stutter (PWS) fall on the lower end of a presumed normal speech motor skill continuum, while persons who do not stutter (PNS) are distributed across the higher end of the continuum. The underlying assumption is that if PWS are at the lower end of a speech motor skill continuum one should be able to find differences between PWS and PNS in tasks which tax their abilities to control their speech motor system. Specifically, these differences should be reflected in factors that typically characterize motor skill, such as the ability to compensate and adapt to changing and/or increasing task demands (Van Lieshout et al., 2004).

A method to increase demands on the speech motor system is through the use of oral-articulatory (bite-block) perturbations (McFarland & Baum, 1995). The presence of a bite-block decreases vocal tract constriction and increases the front cavity cross sectional area resulting in frequency increase in the first formant (F1) and decrease in the second formant (F2) with effects averaging approximately 20 to 100 Hz (McFarland & Baum, 1995). Although, much of the compensatory changes in format frequencies occur early (i.e. immediate compensation), studies on PNS indicate that compensation to bite-block perturbation may develop over time (adaptation) reaching normal acoustic pre-perturbation frequency ranges for vowels after about 10-15 minutes of speaking practice with the bite-block in place and even longer for consonants like fricatives (Baum, McFarland, & Diab, 1996; McFarland & Baum, 1995).

If delayed compensation is found in PNS with perfectly intact sensory-motor systems, then one would expect that if PWS have speech motor skill limitations, they would compensate to a lesser degree (i.e., with higher F1 and lower F2 values and lower perceptual speech quality ratings) and/or may take longer time to adapt to oral articulatory perturbations relative to PNS. For this study, we also included measures of fundamental frequency (F0) of vowels, word and vowel duration. F0 of vowels (especially for high vowel /i/) is related to neural and/or mechanical coupling between oral-articulatory and laryngeal subsystems. Raising the tongue body for high vowels in the presence of a bite-block may result in tensing of the vocal folds and subsequent changes in the F0 of vowels (Whalen & Levitt, 1995). We expect this effect to be stronger for PWS, especially with increasing speech rate, given their presumed limited ability in controlling and coordinating oral and laryngeal subsystems to mitigate the impact of the bite-block. Word and vowel duration measures provide an index of how task requirements relating to speech rate are implemented by PWS and PNS.

2. METHODS

2.1 Subjects

Five male adult PWS (M = 26.1, S.D. = 9.1) and five male adult PNS (M = 25.3, S.D. = 4.2) participated in this study. Based on stuttering severity instrument, two PWS were rated moderate and the remaining three were rated to be very mild, mild and severe. For the perceptual part of the study, listeners included 30 speakers of Canadian English (M= 25.9 yrs., S.D. = 4.7). Subjects reported no problems in vision, hearing and language (other than stuttering for the PWS) prior to participating in this study. The Human Ethics Review Committee at the University of Toronto approved the study.

2.2 Stimuli and Procedures

Subjects reiterated a bi-syllabic nonword /bapi/ for approximately 12 seconds at a pre-determined self-paced speech rate (habitual rate or fast but intelligible). Subjects started with a jaw-free condition (JF), immediately followed by the insertion of the bite-block between the 1st premolars (immediate bite-block [BB] condition) which was then followed by a 10-minute practice period (reading aloud with the bite-block in place), subsequent to which a post-practice (PBB) recording was carried out. For the bite-block conditions the inter-incisor distance was fixed at about 12 mm. Immediate compensatory changes relate to differences between the JF and BB sessions, while delayed adaptation changes are defined as the differences between BB and PBB sessions.

All audio signals were tape recorded and then digitized (16-bit / 16 kHz sampling), using PRAAT (Boersma & Weenink, 2006). For the perceptual part of the study, an AX paradigm was used, where the medial iteration of /bapi/ in the jaw free condition was chosen as the exemplar stimulus (A) and the test stimuli (X) were initial, medial or final iterations of the target word /bapi/ from each session

(JF/BB/PBB). 30 listeners then rated the speech quality of the test stimulus (X) on how closely it resembles the exemplar (A), by moving an indicator on a 10-cm computerized visual analog scale. A rating of 10 implies that the test stimulus is a perfect reproduction of the exemplar and a rating of 0 would mean that the test stimulus is a very poor reproduction of the exemplar. All stimuli in the AX paradigm were blocked by speech rate and speaker.

3. RESULTS AND DISCUSSION

For speech acoustics, first and second (F1 and F2) formants, duration, pitch and intensity values relating to vowels /a/ and /i/ from the bisyllabic nonword (/bapi/) were analyzed. The results of the acoustic evaluation revealed immediate, but partial, vocal tract compensation in the presence of a bite-block. In general, these changes were related to significant increases in F1 for /i/ [F(2,14) = 13.94, p = 0.000] and decreases in F1 and F2 for /a/ in the immediate BB condition. Some of these effects disappeared as speakers adapted to the presence of a bite-block with practice (e.g. F1 for /i/ returned to pre-perturbed JF baseline values following practice). Other spectral effects failed to return to pre-perturbation JF values even after practice (e.g. F2 for /a/) for both groups.

In terms of interactions, for F1 values of /a/ there was a significant two-way group x session interaction [F(2,14) = 5.09, p = 0.021] that was characterized by a significant decrease in F1 values for PWS from JF to BB and PBB conditions, while PNS showed minimal changes across conditions, indicating limited compensation and adaptation abilities in PWS. Furthermore, only for PWS, changes in F1 for /i/ across conditions were mirrored by pitch changes for the vowel as evidenced by a trend for a group x session interaction [F(2,12) = 3.58, p = 0.060]. This suggests a lack of fine-tuning in the control of compensatory responses.

The data also revealed that PWS may require additional time compared to PNS to adapt to perturbations. For example, pitch for /i/ returned to pre-perturbation baseline values for PWS after 10-minutes of practice with the bite-block, but this amount of practice was not sufficient to show a similar change in F1 for /a/. On the other hand, PNS immediately adjusted their vocal tract responses for these variables and these values remained more or less constant across all conditions. PWS also had longer word and vowel durations and spoke louder than PNS, but these factors did not interact with condition (JF/BB/PBB). Moreover, with increases in speech rate the predicted increase in difficulties in compensating and adapting to the bite-block perturbation for PWS was not found.

For the perceptual part of the study, the JF condition had significantly higher speech quality ratings than BB or PBB conditions at both normal [F(2,58) = 178.12, Greenhouse-Geisser adjusted, p = 0.000] and fast rates of speech [F(2,58) = 191.92, Greenhouse-Geisser adjusted, p = 0.000]. The fact that the JF session resulted in the highest quality ratings and did not result in any group effects indicates the reliability of the testing methodology. Importantly, a significant three-way group x condition x rate interaction [F(2,58) = 11.48, p = 0.000] was found. A two-way repeated measures ANOVA was carried out separately for each speech rate, with group (PWS/PNS) and condition (JF/BB/PBB) as within-subject factors. The analysis indicated a group effect only at fast rates of speech [F(1,29) = 48.57, p = 0.000], with significantly higher ratings for PNS in comparison to PWS (see Figure 1).

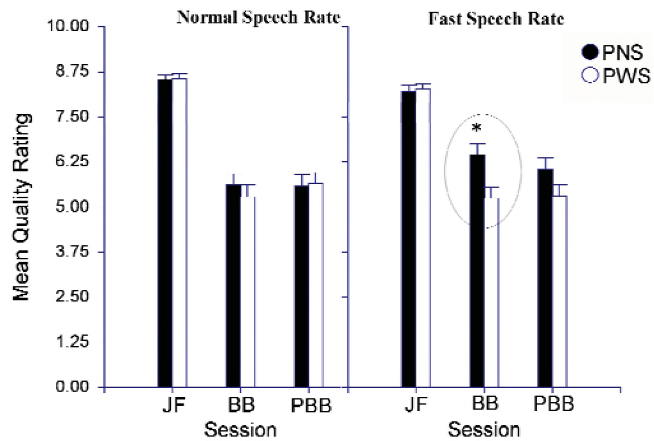

Figure 1. Mean quality rating for normal and fast speech rate.

4. CONCLUSIONS

The findings indicate that PWS have subtle difficulties in compensating and adapting to the presence of a bite-block perturbation not found in PNS. These differences support the claim that PWS may have speech motor skill limitations (Van Lieshout et al., 2004). These findings also highlight the need to assess multiple levels in speech production when looking at compensation and adaptation.

REFERENCES

Baum, S. R., McFarland, D. H., & Diab, M. (1996). Compensation to articulatory perturbation: Perceptual data. *Journal of the Acoustical Society of America, 99,* 3791-3794.

Boersma, P., & Weenink, D. (2006). *Praat: doing phonetics by computer* (Version 4.4.12) URL http://www.praat.org. Viewed 2/26/2008.

McFarland, D. H., & Baum, S. R. (1995). Incomplete compensation to articulatory perturbation, *Journal of the Acoustical Society of America, 97,* 1865-1873.

Van Lieshout, P. H. H. M., Hulstijn, W., & Peters H. F. M (2004). Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In: B. Maassen, R. Kent, H.F.M. Peters, P. Van Lieshout, & W. Hulstijn (eds.). *Speech Motor Control In Normal And Disordered Speech* (pp. 313-355). Oxford, UK: Oxford University Press.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. Journal of Phonetics, 23, 349-366.