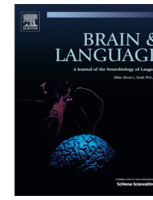
Brain & Language 129 (2014) 24–29

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l



Short Communication

Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter

Shanqing Cai [a,*], Deryk S. Beal [b], Satrajit S. Ghosh [c], Frank H. Guenther [a,d], Joseph S. Perkell [a]

[a] *Department of Speech, Language and Hearing Sciences, Boston University, Boston, MA, USA*
[b] *Institute for Stuttering Treatment and Research and Department of Speech Pathology and Audiology, Faculty of Rehabilitative Medicine, University of Alberta, Edmonton, Alberta, Canada*
[c] *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA*
[d] *Department of Biomedical Engineering, Boston University, Boston, MA, USA*

ARTICLE INFO

Article history:
Accepted 1 January 2014
Available online 31 January 2014

Keywords:
Speech production
Auditory feedback
Stuttering
Sensorimotor integration
Speech timing

ABSTRACT

Auditory feedback (AF), the speech signal received by a speaker's own auditory system, contributes to the online control of speech movements. Recent studies based on AF perturbation provided evidence for abnormalities in the integration of auditory error with ongoing articulation and phonation in persons who stutter (PWS), but stopped short of examining connected speech. This is a crucial limitation considering the importance of sequencing and timing in stuttering. In the current study, we imposed time-varying perturbations on AF while PWS and fluent participants uttered a multisyllabic sentence. Two distinct types of perturbations were used to separately probe the control of the spatial and temporal parameters of articulation. While PWS exhibited only subtle anomalies in the AF-based spatial control, their AF-based fine-tuning of articulatory timing was substantially weaker than normal, especially in early parts of the responses, indicating slowness in the auditory–motor integration for temporal control.

1. Introduction

Auditory feedback (AF) refers to the speech sounds received by the speaker's own auditory system during speech production. AF is an important component of the mechanisms underlying the online control of speech movements. There is evidence (Foundas et al., 2004; Kalinowski, Armson, Roland-Mieszkowski, Stuart, & Gracco, 1993) for abnormalities in the utilization of AF by the speech motor system in stuttering, a developmental disorder of speech fluency in which the production of speech is interrupted by sound or syllable repetitions, prolongations, and silent blocks.

When sudden-onset perturbations are introduced to specific acoustic parameters of AF, normal speakers show online corrections in their production, in directions opposite to the perturbations. Such short-latency (~150 ms) compensatory responses have been demonstrated for fundamental frequency (F0; e.g., Chen, Liu, Xu, & Larson, 2007) and formant frequencies (e.g., Purcell & Munhall, 2006; Tourville, Reilly, & Guenther, 2008), highlighting the active role AF plays in assisting feedforward mechanisms (Guenther, Ghosh, & Tourville, 2006) during online control of phonation and articulation. Recent studies have shown weaker-than-normal compensatory responses to these types of AF perturbation in PWS (for F0, see Loucks, Chon, & Han, 2012; for formant, see Cai et al., 2012). These results indicate that the speech motor system of a PWS cannot compare the desired and actual auditory outcome of speech movements and/or transform the difference (i.e., termed *auditory error*) to proper corrective movements as effectively as non-stutterers can.

How may this subnormal auditory–motor interaction in online speech motor control be manifested during multisyllabic, connected speech? In stuttering, dysfluencies are more likely to occur during multiword utterances than during single words; the frequency of stuttering is positively related to utterance length and complexity (e.g., Soderberg, 1966). Thus examining connected speech production appears to be important for understanding the role of abnormal AF-based speech motor control in this disorder. However, the aforementioned AF perturbation studies (Cai et al. 2012; Loucks et al., 2012) used sustained phonation and isolated monosyllabic words, which were not suitable for probing auditory–motor interaction in stutterers' connected speech.

We have used the technique of time-varying formant perturbation to demonstrate the role of AF in the online control of multisyllabic articulation in normal speakers (Cai, Ghosh, Guenther, & Perkell, 2011). By introducing different types of manipulations of the formant trajectories during the utterance "I owe you a yo-yo", this technique separately examined the spatial and temporal aspects of the control. First, the spatial perturbation altered the formant frequencies at specific local extrema in the AF, while

* Corresponding author. Address: 677 Beacon St, Room 102, Boston, MA 02215, USA. Fax: +1 617 353 7755.
E-mail address: scai@bu.edu (S. Cai).

0093-934X/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandl.2014.01.002

preserving the timing of the extrema. In articulatory terms, this perturbation corresponded to perceived changes in the positions of the articulators (tongue and lips). Under the spatial perturbation, typically fluent participants were shown to compensate by altering formant frequencies produced in the ensuing part of the utterance. Second, the temporal perturbation altered the timing of the formant-frequency extrema in the AF, while preserving the values at those extrema, which corresponded to changes in the timing of the phonemes. Healthy speakers showed timing adjustments in their articulation after the onset of the temporal perturbation and throughout the rest of the utterance.

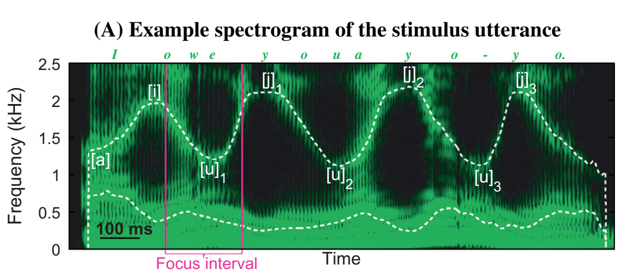
The goal of the current study was to examine whether PWS show deficits in the online AF-based control of multisyllabic articulation using the same technique as Cai et al. (2011). The compensatory responses by a group of PWS to the spatial and temporal perturbations were compared with the responses from fluent controls. Differences in the magnitude and timing of the compensation were analyzed to reveal anomalies in the auditory–motor interaction during multisyllabic articulation in PWS.

2. Results

PWS and matched controls produced the utterance "I owe you a yo-yo". The choice of this utterance was based on the consideration that it consisted of only vowels and semivowels and hence elicited continuous phonation. This allowed us to indirectly measure the spatial positions and timing of the articulation using formant trajectories throughout the utterance.

As Fig. 1 illustrates, there is a set of well-defined local minima and maxima in the second-formant (F2) trajectory, due to lip rounding and the alternation between front and back tongue positions. These extrema were used as landmarks for defining the onsets and offsets of syllables in this utterance, so that we could extract articulatory timing, as well as measure the formant values at the landmarks, which reflect the underlying articulatory positions. Both the spatial and temporal types of AF perturbation occurred during the word "owe" and the transition from "owe" to the following word "you", as indicated by the focus interval in Fig. 1A. As an initial part of each experiment, the participant was trained to produce the sentence within medium ranges of speech intensity (74–84 dB SPL at 10 cm from mouth) and speaking rate (sentence duration: 1.0–1.4 s).

We conducted two experiments on a group of adults with persistent developmental stuttering confirmed by a certified speech-language pathologist (D.S.B.), in addition to two different groups of persons with fluent speech (PFS) as matched controls. Experiment 1 focused on perturbations of spatial parameters in the F2 trajectory; Experiment 2 used perturbations of temporal parameters. Each PWS undertook both Experiments 1 and 2, in randomized order. Two different but partially overlapping groups of controls participated in Experiments 1 and 2. In the following, we visit the results from the spatial perturbations in Experiment 1, then we examine the results from the temporal perturbations in Experiment 2.



(B) List of *ad hoc* phonetic symbols used in the current paper

[a]	Onset of "I"	$[u]_2$	F2 minimum at the end of "you"
[i]	F2 maximum at the end of "I"	$[j]_2$	F2 maximum at the onset of the 1st "yo"
$[u]_1$	F2 minimum at the end of "owe"	$[u]_3$	F2 minimum at the end of the 1st "yo"
$[j]_1$	F2 maximum at the onset of "you"	$[j]_3$	F2 maximum at the onset of the 2nd "yo"

Fig. 1. An example spectrogram of the stimulus utterance "I owe you a yo-yo", with the F1 and F2 trajectories overlaid (Panel A, dashed curves). The set of local minima and maxima in F2 (landmarks) are labeled by the phonetic symbols (Panel B). The focus interval is the time period containing the AF perturbation.

2.1. Experiment 1: Spatial perturbation

Twenty PWS and 37 PFS participated in Experiment 1, which focused on the AF-based control of the spatial parameters of multisyllabic articulation, as reflected in formant values. The age distributions of the two groups were similar (mean ± 1 SD: PWS 27.0 ± 7.7; PFS: 24.9 ± 5.6; *t*-test: $p = 0.24$); so were the gender distributions (PWS: 4F16 M; PFS: 6F31 M; χ^2-test: $p = 0.94$). The stuttering severity of the PWS participants, as measured with Stuttering Severity Instrument version 4 (Riley, 2008), ranged from 13 to 43 (median: 25.4; interquartile range: 11.5).

As the examples in Fig. 2A illustrates, the Up perturbation increased the value of F2 at the local minimum corresponding to the end of the word "owe", in a way that preserved the smoothness of the F2 trajectory. The Down perturbation decreased the F2 at the local minimum (Fig. 2B). Such changes in the F2 value would result naturally from changes in the front-back position of the tongue and/or the degree of lip rounding during the [u] sound in "owe". Both the Up and Down perturbations preserved the timing of the local F2 minimum. Therefore they focused on altering the acoustic correlates of the spatial parameters of articulation.

To analyze the compensatory changes in the F2 values produced by the participants under the perturbations, we manually extracted the seven local extrema ([i] to $[j]_3$ as listed in Fig. 1B) as landmark points from each trial. We manipulated the time axis in each trial in a piece-wise linear fashion, so as to align all trials at these landmarks. Specifically, the time between each pair of adjacent landmarks were linearly interpolated at 250 evenly spaced points, giving rise to a single piecewise-normalized time axis (e.g., Fig. 2C–E) on which the F2 values were analyzed. This time normalization followed the approach of Cai et al. (2011). Comparisons between the perturbation conditions and between the groups were performed on this piecewise-normalized time axis using Monte Carlo permutation tests (see Methods).

The PFS responded to the Down and Up perturbations by altering the F2 values in their production in directions opposite to the perturbations (Fig. 2C and D: black curves). Under the Up perturbation, the earliest significant compensation could be seen between $[u]_1$ and $[j]_1$ (i.e., during the transition from "owe" to "you"). Under the Down perturbation, a significant response (corrected) started shortly after $[u]_1$ (the end of "owe") and exhibited multiple local maxima between $[u]_1$ and $[j]_1$, between $[j]_1$ and $[u]_2$, and between $[u]_2$ and $[j]_2$. The contrast between the Down and Up F2 trajectories (black curve in Fig. 2E) showed a similar pattern, with the significant compensation seen as early as between [i] and $[u]_1$ (i.e., during "owe") and as late as between $[u]_2$ and $[j]_2$ (i.e., after "you"). In general, the compensatory responses were longer and slightly greater in magnitude under the Down perturbation than under the Up one. This counteracting and slightly asymmetric pattern of response is highly similar to the results in Cai et al. (2011).

As shown by the red curves in Fig. 2C and D, the mean responses to the spatial perturbations in PWS group were similar to those from the PFS group in that they opposed the directions of perturbation. However, compared to the PFS, trends of later response onset and slower ramping to peak response can be seen PWS group (Fig. 2C and D). Under the Up perturbation, the peak compensation