ABSTRACT

The Role of Auditory Feedback in the Online Control of Articulatory Trajectories and Timing in a Multi-syllabic Utterance

S. Cai, S.S. Ghosh, J.S. Perkell, and F.H. Guenther

Abstract (198 words; 200 word limit)

In order to study the role of auditory feedback in the online control of articulatory timing and trajectories in multi-syllabic utterances, we imposed unexpected time-varying perturbations to the 2nd formant frequency trajectory when subjects produce the utterance "I owe you a yo-yo". The perturbations either accelerated or decelerated the F2 transition between /i/ in the first word "I" and /u/ in the second word "owe", called *acceleration-perturbation* (AP) and *deceleration-perturbation* (DP), respectively. Here we report preliminary results from 12 subjects.

Compared to the unperturbed (control) trials, the AP trials showed a shortening in the durations of the word "owe" and earlier initiation of the word "you" following the perturbation, whereas the DP trials showed lengthened durations in "owe" and delayed the initiations of "you". The timing differences between the AP and DP conditions each reached statistical significance. In addition to the timing changes, perturbed trials also showed significant changes in the path (magnitude) of F2 compared to the unperturbed baseline. These F2 path changes were in the directions opposite to the perturbations. These observations indicate that both the timing and path of articulation can be fine-tuned online by utilizing auditory feedback with a short latency during multisyllabic utterances.

PROPOSAL SUMMARY

The Role of Auditory Feedback in the Online Control of Articulatory Trajectory and Timing in a Multi-syllabic Utterance

Summary Text (750 words; 750 word limit)

Introduction

Speakers compensate for unexpected perturbations to auditory feedback by altering their productions in near-real time ([1,2]), indicating a role of auditory feedback in online control of articulation. However, those studies used prolonged quasi-steady-state vowels, hence it is still unclear whether auditory feedback is involved in the online control of articulation during multi-syllabic, time-varying utterances at normal speaking rates. In this study, we aim to address this question by examining the changes in syllabic timing and articulatory trajectories caused by perturbations to the second formant frequency (F2) during production of the utterance "I owe you a yo-yo".

Methods

Twelve healthy native-English-speaking subjects participated in this study. The subjects were trained to produce the utterance within durations between 1.2 and 1.6 seconds (i.e., approximately normal speaking rates).

The all-vocalic utterance "I owe you a yo-yo" was chosen to increase the extent to which the articulatory movements are reflected in acoustics. A computer program (based on [3]) was used to detect the segment between the F2 maxima corresponding to /i/ in "I" and /j/ in "you" and to impose F2 perturbations on this segment online (11-ms delay). Two types of perturbation were given: (1) acceleration-perturbation (AP), in which the speed of the F2 movement was increased by approximately 25% (e.g., Fig. 1B), and (2) deceleration-perturbation (DP), in which the speed of the F2 movement was decreased by 25% (Fig. 1C).

An experiment contained 120 repetitions of the utterance, arranged in 20 blocks of 6. In each block, there was one AP trial, one DP trial, and four unperturbed control trials. Order of the trials was randomized.

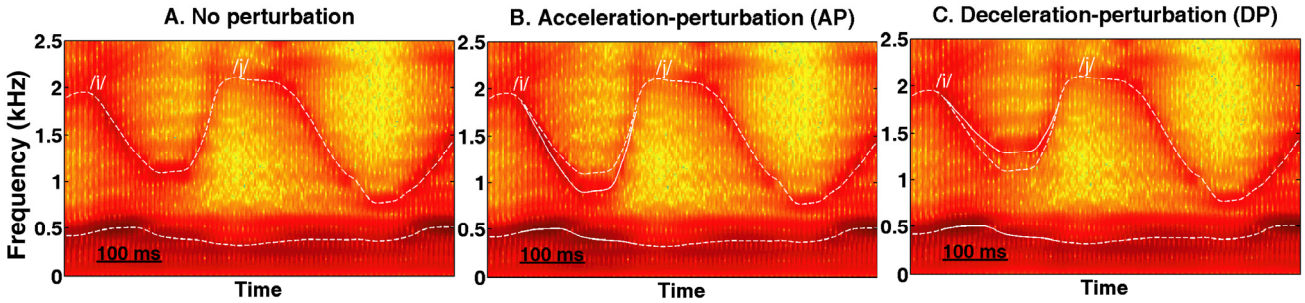

Figure 1. A: original spectrogram. **B**: AP of the signal in A. **C**: DP of the signal in A. Dashed white traces: original F1 and F2; solid white traces: perturbed F2.

Results and Discussion

The average changes in the interval between the F2 maximum corresponding to /i/ in "I" and the F2 minimum corresponding to /u/ in "you" (called /i/-/u/$_1$ interval) re control trials are shown in Fig. 2B. It can be seen than the AP and DP trials showed decreased and increased /i/-/u/$_1$ intervals, indicating earlier- and later-than-control completion of the syllable "owe", respectively. The contrast between AP and DP reached statistical significance (RM-ANOVA; $F(1,11)=5.06$, $P<0.05$). The interval between the F2 maxima corresponding to /i/ in "I" and /j/ in "you" showed similar changes, indicating that AP and DP led to earlier and delayed initiation of the word "you", respectively (Fig. 2C; non-significant trend: $F(1,11)=2.64$, $P=0.13$).

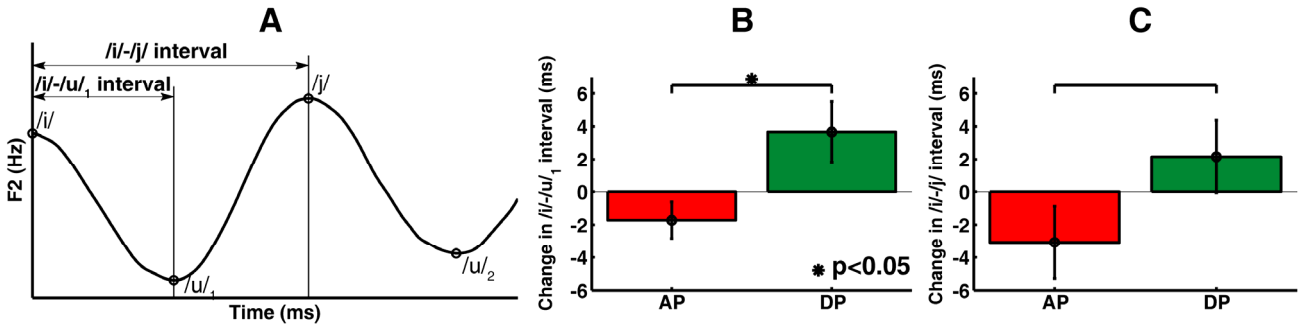

Figure 2. Perturbation-induced timing changes. **A**: schematic drawing showing the two time intervals. **B**: the group average changes (re control) in /i/-/u/$_1$ interval in the AP and DP trials. Error bars: ±1 SEM. **C**: changes in /i/-/j/ interval in the AP and DP trials.

In addition to the timing changes, the perturbed trials also showed significant changes in the magnitude of F2. Fig. 3A shows the time-normalized group-average F2 trajectories between the F2 maxima corresponding to /i/ in "I" and the F2 minimum corresponding to /u/ in "you" (termed /u/$_2$). The time normalization involved anchoring the F2 traces from individual trials at four anchor points: /i/, /u/$_1$, /j/ and /u/$_2$ (Fig. 3A), and linear time interpolation between the anchors. Analysis was done on the anchors (T1, T3, T5, and T7) and the mid-points (T2, T4, and T6).

As Fig. 3B and C show, there was no significant difference between the AP and DP trials at T1 and T2 (i.e., /i/ and the midpoint between /i/ and /u/$_1$), which is not surprising given close proximity between these two time points and the perturbation onset (<~60 ms). Significant differences were observed at T3 (/u/$_1$), T4 (midpoint between /u/$_1$ and /j/), and T6 (midpoint between /j/ and /u/$_2$) (Fig. 3D-G). Time point T5 (/j/) showed the same trend as T4 but didn't reach statistical significance due to the small magnitude of difference, which may be attributable to the quantal articulatory-acoustical relation near /j/ ([4]). In general, the directions of these changes at T3 – T7 were opposite to the directions of the perturbations, indicating compensatory corrections made to the articulatory trajectory online.

To our knowledge, this is the first demonstration of the role of auditory feedback in online control of timing and magnitude of articulation in multi-syllabic utterances at close-to-normal speaking rates.

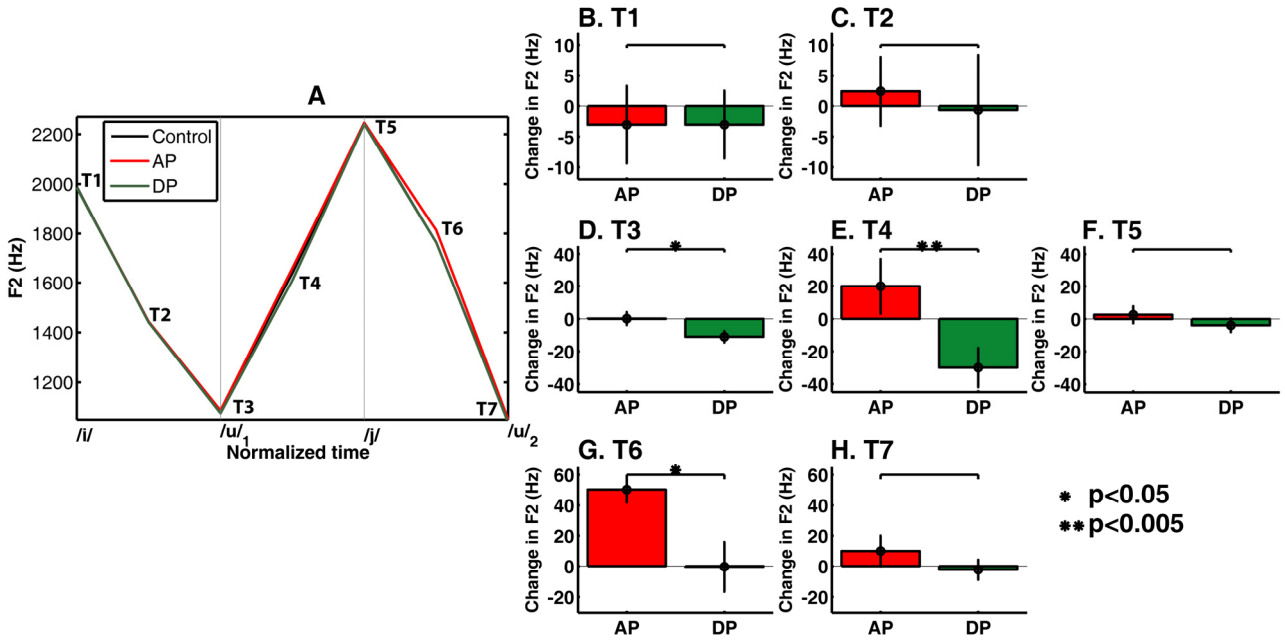

Figure 3. Perturbations-induced changes in F2 magnitudes. **A**: The time-normalized average F2s at four anchor points (T1,T3,T5,T7) and three midpoints (T2,T4,T6). B-H: The changes in the AP and DP trials re control at the seven normalized times. Error bars: ±1 SEM.

Acknowledgements

This work was supported by NIH grant R01DC01925.

REFERENCES

1. Purcell DW, Munahll KG. (2006a). "Compensation following real-time manipulation of formants in isolated vowels." *J. Acoust. Soc. Am.*, 119(4):2288-2297.
2. Tourville JA, Reilly KJ, Guenther FH. (2008). "Neural mechanisms underlying auditory feedback control of speech." *NeuroImage*, 39:1429-1443.
3. Cai S, Boucek M, Ghosh SS, Guenther FH, Perkell JS. (2008). "A system for online dynamic perturbation of formant frequencies and results from perturbation of the Mandarin triphthong /iau/." In *Proceedings of the 8th Intl. Seminar on Speech Production*, Strasbourg, France, December, Dec. 8-12, 2008. pp. 65-68.
4. Stevens KN. (1989). "On the quantal nature of speech." *J. Phonetics.*, 17:3-45.