Title
Phase-resetting and rhythmic pattern generation in speech production

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Author
Saltzman, Elliot

Publication Date
1996

Peer reviewed

Phase-resetting and rhythmic pattern generation in speech production

Elliot Saltzman
Haskins Laboratories, 270 Crown Street, New Haven, CT;
Department of Psychology, University of Connecticut, Storrs, CT
saltzman@haskins.yale.edu

The postulation of segmental or phonemic units in linguistics implies that there is some degree of cohesion among the articulations that comprise these segments. That is, we expect that the cohesion among *gestures* within segmental units is stronger, in some sense, than the cohesion among gestures of different segmental units. (The term gesture is used here to denote a member of an equivalence class of articulatory movement patterns that are *actively* controlled with reference to a given speech-relevant goal, e.g., coordinated movements of the lips and jaw to produce a bilabial closure for a /p/). However, the nature and origin of this intergestural "glue" are issues that require empirical study. It has been hypothesized from a dynamical systems approach that intergestural cohesion can be accounted for by coupling structures defined among intrasegmental gestural units (Saltzman & Munhall, 1989). If so, evidence of such coupling should be experimentally observable.

We report results from a series of phase-resetting (e.g., Winfree, 1980) studies of speech production indicating that intergestural temporal cohesion is greater within segments than between segments. In our data, the coordinated gestures are a bilabial closing and a laryngeal devoicing gesture for a single /p/, as compared to those for /p/s at the beginnings of successive syllables. In these studies, downward-directed mechanical perturbations are applied to the lower lip during the repetitive, rhythmic utterance /...paepaepae.../ and the nonsense word "puh'saepaepple" (e.g, Saltzman, Löfqvist, K.-Shaw, Kay, & Rubin, 1995).

In the repetitive utterance, the amount of temporal shift introduced by a perturbation is measured relative to the timing pattern that existed prior to the perturbation. This shift is measured after the transient, perturbation-induced distortions to the rhythm have subsided, and the system has returned to its pre-perturbation, steady-state rhythm. The finding of a post-perturbation, steady-state temporal shift using this method supports the hypothesis that there exists a central "clock" that drives the articulatory periphery and whose state is altered (phase-shifted) by feedback specific to events at the periphery. Perturbations that induce transient temporal shifts, but not steady-state shifts, do not indicate resetting of the central clock. A further, crucial aspect of the phase-resetting paradigm is that, across trials, perturbations are delivered so as to sample all phases of the rhythm's cycle in order to examine the variation, over the course of the cycle, in the sensitivity of the central clock to peripheral events.

Using this paradigm, we showed that perturbations delivered during the speech sequence /...paepaepae.../ induce systematic steady-state shifts in the timing between the bilabial closing and laryngeal devoicing gestures for /p/s at the beginnings of successive syllables, and observed smaller steady-state shifts in the *relative* phasing of these gestures within the /p/s. Thus, these results not only demonstrate a resetting of the central "clock" for these utterances, but also demonstrate that intergestural temporal cohesion is greater within segments than between segments. That is, the individual temporal shifts of the bilabial and laryngeal gestures are large compared to the relative temporal shift between these gestures, and the lips and larynx appear to be phase-advanced as a relatively coherent unit. Furthermore, such resetting behavior occurs only when the perturbation is delivered within a "sensitive phase" of the cycle. During this period, the downwardly directed lower lip perturbation opposes the just-initiated, actively controlled bilabial closing gesture for /p/. Thus, the sensitive period corresponds (roughly) to the acceleration portion of the closing gesture (Kawato, personal communication). Additionally, although changes in temporal structure were found for other perturbed phases, these changes were simply transient effects, and do not indicate a resetting of the central "clock." Finally, in conjunction with the repetitive utterance, it is important to analyze a corresponding nonrepetitive utterance that contains the same target sequence. By comparing the transient changes in speech timing induced by perturbations for the repetitive and nonsense word utterances, we conclude that the behaviors observed in the (relatively unnatural) rhythmic sequences and the (more natural) nonsense word sequences are governed by a common set of dynamical principles.

These phase-resetting results imply that each gesture in a rhythmic utterance is governed by a corresponding oscillatory unit (or set of units), and that during the performance of a given sequence these oscillators behave as functionally coupled, nonlinear, limit-cycle oscillators. In unperturbed cases, the observed pattern of gestural activity corresponds to an associated pattern of synchronization (entrainment) and relative phasing among the oscillators that is specific to the utterance. When the system is perturbed in the rhythm's sensitive phase, the entire rhythm is phase-advanced in the steady-state and the relative phasing among gestural units is altered to a smaller degree. Such steady-state shifts of intergestural relative phasing will be seen if either: a) the experimental observation time (approximately 20 syllables) was shorter than the *relaxation time* required to return to the system's pre-perturbation relative phase value; or b) the initially observed relative phasing was simply one value in a *phase window* (Byrd, in press) or interval of allowable relative phases.

This interpretation is also consistent with experiments in which discontinuous transitions of intergestural phasing accompanied continuous increases in speaking rate (e.g., Kelso, Saltzman, & Tuller, 1986a, 1986b; Tuller & Kelso, 1991). Specifically, when subjects spoke the syllable /pi/

repetitively at increasing rates, the relative phasing of the bilabial and laryngeal gestures associated with the /p/ did not change from the pattern observed at a self-selected, comfortable rate. However, when the repeated syllable /ip/ was similarly increased in rate, its relative phasing pattern switched relatively abruptly at a critical speed—from that observed for a self-selected, comfortable rate to the pattern observed for the /pi/ sequences. Such intergestural *phase transitions* may be viewed as behaviors of a system of nonlinearly coupled, limit-cycle oscillators that bifurcate from one modal pattern that becomes unstable with increasing rate to another modal pattern that retains its stability (e.g., Haken, Kelso, & Bunz, 1985).

Acknowledgement

Work described in this paper was supported by NIH Grant DC-00121 (Dynamics of Speech Articulation).

References

Byrd, D. (in press). A phase window framework for articulatory timing. *Phonology*.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986a). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics, 14*, 29-60.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. (1986b). Intentional contents, communicative context, and task dynamics: A reply to the commentators. *Journal of Phonetics, 14*, 171-196.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*, 333-382.

Saltzman, E., Lofqvist, A., Kinsella-Shaw, J., Kay, B., & Rubin, P. (1995). On the dynamics of temporal patterning in speech. In F. Bell-Berti, & L. Raphael, (Eds.). *Studies in speech production: A Festschrift for Katherine Safford Harris*. Woodbury, New York: American Institute of Physics.

Tuller, B. & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research, 34*, 501-508.

Winfree, A. T. (1980). *The geometry of biological time*. NY: Springer-Verlag.