Exp Brain Res (1998) 123:412–424

RESEARCH ARTICLE

1105

Elliot Saltzman · Anders Löfqvist · Bruce Kay
Jeff Kinsella-Shaw · Philip Rubin

Dynamics of intergestural timing: a perturbation study of lip-larynx coordination

Received: 21 March 1997 / Accepted: 9 June 1998

Abstract In this study, downward-directed mechanical perturbations were applied to the lower lip during both repetitive (/...pæpæpæ.../) and discrete (/pə'sæpæpl/) utterances in order to examine the perturbation-induced changes of intergestural timing between syllables (i.e., between the bilabial and laryngeal gestures for successive /p/'s) and within phonemes (i.e., between the bilabial and laryngeal gestures within single /p/'s). Our findings led us to several conclusions. First, steady-state (phase-resetting) analyses of the repetitive utterances indicated both that "permanent" phase shifts existed for both the lips and the larynx after the system returned to its pre-perturbation rhythm and that smaller steady-state shifts occurred in the relative phasing of these gestures. These results support the hypothesis that central intergestural dynamics can be reset by peripheral articulatory events. Such resetting was strongest when the perturbation was delivered within a "sensitive phase" of the cycle, during which the downwardly directed lower-lip perturbation opposed the just-initiated, actively controlled bilabial closing gesture for /p/. Although changes in syllable duration were found for other perturbed phases, these changes were simply transient effects and did not indicate a resetting of the central "clock." Second, analyses of the transient portions of the perturbed cycles of the repetitive utterances indicated that the perturbation-induced steady-state phase shifts are almost totally attributable to changes occurring during the first two perturbed cycles. Finally, the transient changes in speech timing induced by perturbations in the discrete sequences appeared to share a common dynamical basis with the changes to the repetitive sequences. We conclude by speculating on the type of dynamical system that could generate these temporal patterns.

Key words Speech · Motor control · Phase resetting · Timing · Dynamics

E. Saltzman (✉) · A. Löfqvist · P. Rubin
Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, USA
Fax: +1-203-865-8963

E. Saltzman
Department of Physical Therapy, Boston University, Boston, Massachusetts, USA

B. Kay
Department of Cognitive and Linguistic Sciences, Brown University, Providence, Rhode Island, USA

J. Kinsella-Shaw
University of Connecticut, Storrs, Connecticut, USA

Introduction

Speech production requires the control and coordination of several different systems: the respiratory system, the larynx, and the upper articulators (i.e., the lips, the tongue, the jaw, and the velum). During speech, these systems are coupled in a functional manner to produce the acoustic signal and visual information used for transmitting linguistic intent. This study investigates the principles governing the temporal control of speech production with particular emphasis on intergestural timing in the production of voiceless consonants. For example, the production of the bilabial, voiceless consonant /p/ requires the following set of gestures. The lips are closed by joint activity of the jaw and the upper and lower lip. The velum is elevated to seal off the entrance to the nasal cavity. The glottis is widened momentarily, and the longitudinal tension of the vocal folds is often concomitantly increased to prevent glottal vibrations. These oral and laryngeal actions all contribute to the period of silence in the acoustic signal and the increase in oral air pressure associated with the stop consonant. In addition, the oral closing and opening gestures and the laryngeal abduction/adduction movements have to be coordinated in time to produce the proper acoustic signal. Variations in the timing of these articulatory events are commonly used in different languages to produce linguistically meaningful contrasts of voicing and aspiration (Abramson 1977; Löfqvist 1980, 1992; Löfqvist and Yoshioka 1981).

One experimental approach, which has been used to examine intergestural timing, involves applying a me-

chanical perturbation to one articulator, commonly the jaw or the lower lip, and observing the changes in temporal patterning induced by the load. The results of studies using this experimental technique show that the timing between gestures is affected by such perturbations. For example, Gracco and Abbs (1989) used such an experimental paradigm to examine the sequencing of the oral-closing movements for the bilabial stops in two successive syllables. Their results showed that the interval between the two gestures was affected systematically by the perturbation. Perturbation-induced changes of intergestural timing have also been demonstrated for the gestural components of single phonemes. For example, Munhall et al. (1994) applied a load to the lower lip while subjects were making the closure for the bilabial voiceless consonant /p/. As a result of the perturbation, the onset of glottal abduction for the consonant was delayed, possibly to maintain the phasing between the oral and laryngeal movements at the transition from the vowel to the onset of the stop consonant. However, in the perturbed productions, the duration of the oral closure decreased, while the duration of the glottal abduction/adduction cycle increased. As a consequence, the normal phasing between the oral and laryngeal articulations was disrupted at the release of the oral closure for the stop. Finally, it should be noted that temporal changes in response to perturbations are not restricted to speech motor activities, but have also been demonstrated in the timing of non-speech movements. For example, Kay et al. (1991) showed that transient mechanical perturbations delivered to the upper limbs during unimanual rhythmic tasks can alter the underlying timing structure of the ongoing sequence and induce systematic shifts in the timing of subsequent movement elements. In particular, Kay et al. (1991) observed that their perturbation produced an overall phase advance, which was modulated by where the perturbation occurred in the movement cycle.

The present study uses *phase-resetting* techniques to investigate the temporal control of successive opening and closing movements of the lips and the larynx in voiceless consonant production. In addition, possible transient effects of perturbations on the relative phasing between the lips and the larynx are examined. The goal of phase-resetting analyses (e.g., Glass and Mackey 1988; Guevara et al. 1981; Kawato 1981; Pavlidis 1973; Winfree 1980) is to determine whether perturbations delivered during an ongoing rhythm have a permanent effect (i.e., phase shift) on the underlying temporal organization of the rhythm. Phase-resetting techniques have been used in many kinematic and neurophysiological studies of the control and coordination of rhythmic movements (e.g., Currie and Stein 1989; Lee and Stein 1981; Lennard 1985; Lennard and Hermanson 1985). What is measured in such studies is the amount of temporal shift introduced by the perturbation, relative to the sequence's timing prior to the perturbation. This phase shift is measured after the perturbation-induced transients have subsided and the system has returned to its pre-perturbation, steady-state rhythm. If the perturbation induces such a shift in an extended, repetitive speech sequence, this result would imply that patterns of intergestural relative phasing are not rigidly specified over the sequence. Rather, such results would suggest that gestural patterning evolves fluidly and flexibly over the course of an ongoing sequence, governed by an intergestural dynamical system (i.e., a central timing network or "clock"), which does not simply drive the articulatory periphery in a strictly feedforward, unidirectionally coupled manner. Instead, central and peripheral dynamics must be coupled bidirectionally, so that feedback information from the articulatory periphery can influence the state of the central clock.

It is important to note that relatively lengthy, repetitive utterances are required for the use of the steady-state, phase-resetting technique in order to be able to distinguish temporal articulatory distortions, which are attributable to central resetting processes, from those attributable to the systematic, yet transient behavior of the articulatory periphery. The utterances must be repetitive, since the units of analysis are cycles, and, by definition, successive cycles must be approximately identical. The utterance must be relatively long since, even in the minimum cycle-period case where each cycle is only one syllable long, one needs: (1) a steady-state measure of pre-perturbation behavior that includes approximately 5–10 cycles; (2) 1–2 (occasionally 3) more cycles, during which the perturbation is applied, at least in our previous experiments (Saltzman 1992) and in the experiment reported below; (3) several more cycles (occasionally none) in order to settle back to, within a criterion degree of closeness, the pre-perturbation behavior; and (4) approximately 2–10 cycles to provide a steady-state measure of post-perturbation behavior.

It is also important to note that, of course, normal speech does not consist of extended, rhythmic repetitions of a single syllable. Therefore, in order to be sure that the central phase shifts identified using phase-resetting techniques actually reflect processes governing normal utterances, it is necessary to bridge the theoretical gap between phase-resetting results and those obtained from perturbing discrete, word-like sequences. Because of the relatively short duration of such discrete sequences, the system cannot be relied upon to settle down and "shake off" the effects of the perturbation in the time between the offset of the perturbation and the end of the utterance. In effect, one can reliably study only the transient responses to perturbations in such sequences. Thus, in order to relate steady-state, phase-resetting data meaningfully to transient data obtained by perturbing discrete utterances, it is necessary to study the transient responses of *repetitive* as well as discrete utterances, preferably using a within-subject experimental design. Once the relation between the steady-state and transient patterns is understood for the repetitive data, a conceptual link can be forged between the transient patterns of the repetitive and discrete data, and shared dynamic principles governing articulatory behavior can be identified. Preliminary results from the present study have been described previously (Saltzman et al. 1992, 1995).

Materials and methods

Subjects

Two males (the first two authors of this paper), with no history of language impairment, were subjects; one a native speaker of American English (ES) and the other a native Swedish speaker (AL) fluent in American English. This study was approved by the appropriate ethics committee, and both subjects gave their informed consent prior to their inclusion in this study.

Equipment and data processing

The subjects sat in an adjustable dental chair with the head restrained in an external frame (see Fig. 1). A small paddle connected to a torque motor was placed on the lower lip with a tracking force of 3 g in order to deliver step pulses of downward force (50 g) at random times during the experimental trials. Timing of perturbation onset was controlled by a computer. Oral articulatory movements were measured optoelectronically using infrared light-emitting diodes (LEDs) mounted on the upper lip, lower lip, lip paddle, nose (the nose LED acted as a spatial reference), and a custom-made jaw splint. Laryngeal abduction and adduction movements were recorded using a transillumination technique, in which a fiberoptic endoscope was introduced through the nose and placed in the pharynx in order to illuminate the larynx. The amount of light passing through the glottis, which depends on the degree of laryngeal opening (Baer et al. 1983; Löfqvist and Yoshioka 1980), was detected by an optical sensor placed on a neck collar just below the cricoid cartilage. During the experimental session, the illuminated larynx was displayed on a video monitor to ensure that the view of the larynx was unobstructed and that the endoscope's lens was not fogged. The acoustic speech signal and control voltage applied to the torque motor were recorded. All data were fed into a 16-track FM tape recorder for later digitization and signal processing.

For processing, the audio signal was sampled at 10 kHz, while all movement signals were sampled at 500 Hz. All movement signals, except subject ES's laryngeal trajectory, were smoothed with a 42-ms triangular window; ES's laryngeal trajectory was smoothed with a 82-ms triangular window, since this signal was relatively noisy. After smoothing, the velocity of upper-lip movement was obtained for use in the discrete-trials analyses (see the data-analysis-procedures section below entitled "Discrete sequences") using a three-point central-difference algorithm; the obtained velocity signal was smoothed once more using a 42-ms triangular window.

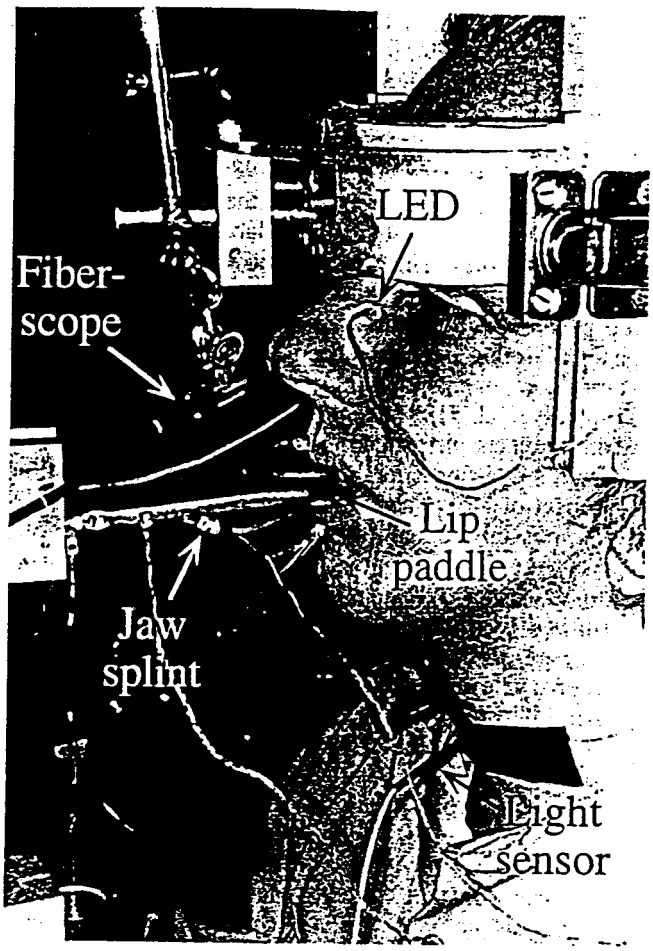

Fig. 1 The experimental setup. Light-emitting diodes (*LED*'s) were attached to the articulators to be tracked. A fiberscope provided light for the transillumination used to record laryngeal behavior. Perturbations were delivered to the lower lip using a lip paddle connected to a torque motor

Protocol

Twelve blocks of 25 trials were performed during each of two sessions. Each session lasted approximately 3 h. Blocks alternated between *repetitive* and *discrete* experimental conditions. In the discrete condition, each trial consisted of the sequence /pə'sæpæpl/; in the repetitive condition, each trial consisted of a sequence of approximately 20–30 repetitions of the syllable /pæ/, spoken at a syllable rate comparable to that used in the discrete trials. Perturbations were delivered during random sampling of 80% of the trials. The subjects were instructed to not actively resist the perturbation and to continue speaking as normally as possible.

For the repetitive blocks, perturbation duration was preset in an external timing circuit to equal the subject's average syllable duration measured from lower-lip trajectories during pre-test repetitive trials. Pre-test measures were also used to parameterize a random timing circuit for controlling perturbation onset. This circuit was triggered by the acoustic release burst of the initial /p/ for both repetitive and discrete sequences. On each perturbed repetitive trial, the perturbation was delivered during the *n*th syllable (*n* varied randomly from 8 to 11), and after *m*% of the predetermined syllable duration (*m* varied randomly from 1 to 100). For the discrete blocks, the duration of the perturbation was preset to equal the subject's average interval between maximum lower-lip lowerings for the first and second /æ/, measured during a set of pre-test productions. On each perturbed discrete trial, perturbation offsets occurred at *x*% of the pre-test interval between maximum lower-lip lowerings for the first and second /æ/ (*x* varied randomly from 1 to 100).

Data-analysis procedures

Repetitive sequences: phase-resetting (steady-state) analyses

Bilabial movements were analyzed using a lip-aperture trajectory, which was defined by subtracting the lower-lip signal from the upper-lip signal (Fig. 2, top panel); laryngeal movements were analyzed using a glottal-opening trajectory, which was defined using the transillumination signal (Fig. 2, middle panel). Individual cycles were then defined between successive peak openings, and four cycle types were identified:

1. *Pre-perturbation* cycles included the trial's first cycle through the last cycle before perturbation onset (see also Fig. 3).
2. *Perturbation* cycles included all cycles that overlapped the perturbation interval (see also Fig. 3).
3. *Transient* cycles were defined as those cycles following the perturbation and during which cycle periods deviated from the average pre-perturbation cycle period by more than approximately 2.5 standard deviations; and

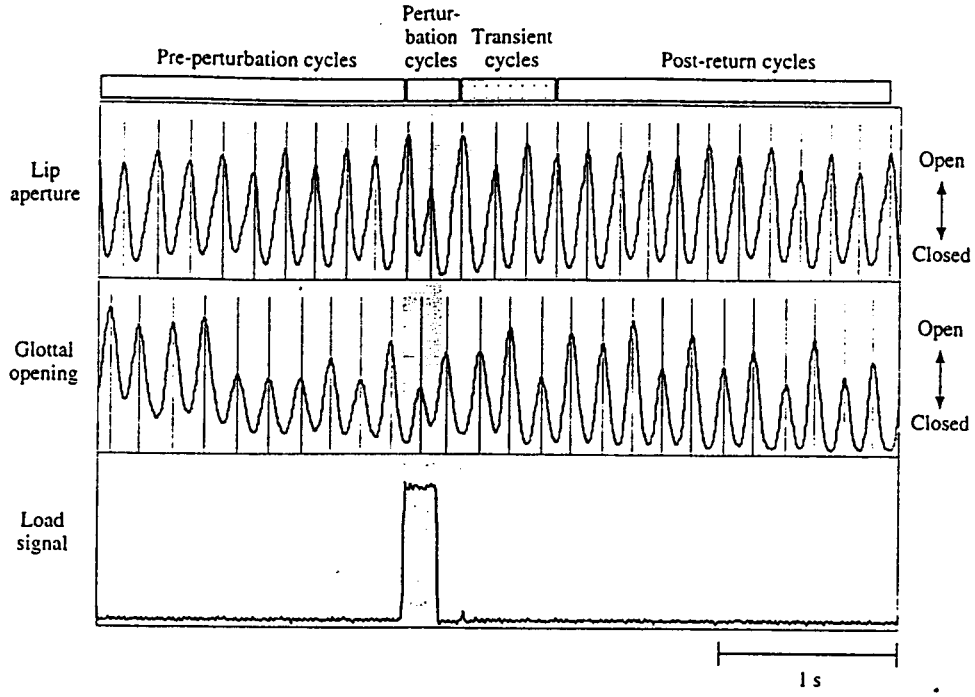

Fig. 2 Data trajectories for a single repetitive trial: lip aperture (*top panel*), glottal opening (*middle panel*), and torque-load signal (*bottom panel*). The sign of the loading signal is inverted in the figure to emphasize that the downward force on the lower lip acted to increase lip aperture. The *boxes* above the top panel mark the different cycle types for the trial

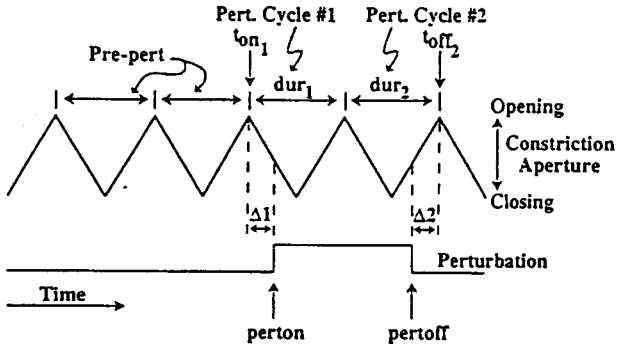

Fig. 3 Schematic display of constriction (bilabial or glottal) aperture and perturbation trajectories for pre-perturbation cycle (*Prepert*) and for perturbation cycles 1 (*Pert. Cycle #1*) and 2 (*Pert. Cycle #2*). t_{on1} Onset time of the first perturbation cycle, t_{off2} offset time of the second perturbation cycle, $\Delta 1$ difference perton–t_{on1}, $\Delta 2$ difference pertoff–t_{off2}, dur_1 duration of perturbation cycle 1, dur_2 duration of perturbation cycle 2

4. *Post-return* cycles were defined from the last transient cycle to the end of the trial.

For any given trial, analyses were limited to a maximum of 20 cycles, approximately the most that could be comfortably produced in one breath.

For both lip-aperture and glottal-opening trajectories, cycle phase, ϕ, was defined to be zero at all peak openings. For all other points between peak openings, phase was defined as (t/T_i), where t is the time (in s) from the most recent peak preceding a given event of interest, and T_i is the period (in s) of the cycle containing the event. Thus, phase values of events occurring in a given cycle were defined in a normalized range from zero to one. In the perturbation trials, the offset time of the perturbation served as a temporal anchoring point for sampling ("strobing") both backward and forward in time into the pre-perturbation and post-return cycle sequences, respectively, using the average pre-perturbation cycle period to define the strobe period. The within-cycle strobe phases from the pre-perturbation and post-return cycles were then averaged to define, respectively, an *average old phase*, $\overline{\phi_{old}}$ (modulo 1), and *average new phase*, $\overline{\phi_{new}}$ (modulo 1). In order to compensate for phase drift between the pre-perturbation and post-return cycles, which was due to any within-trial differences in average periods of the pre-perturbation cycles (which defined the strobe period) and post-return cycles (which were strobed to define the trial's set of new phases), the following correction formula was applied to each trial's *average new-phase* value (see Appendix for derivation):

$$\overline{\phi_{new,corrected}} = \left(\overline{\phi_{new}} - \left(\frac{n-1}{2}\right)\left(\frac{\overline{T_{pre}}}{\overline{T_{post}}} - 1\right)\right) \text{(modulo 1)}. \quad (1)$$

where n = number of post-return cycles; $\overline{T_{pre}}$ = average pre-perturbation cycle period; and $\overline{T_{post}}$ = average post-return cycle period. Without this correction, the mean post-return phase would be overestimated if $\overline{T_{pre}} > \overline{T_{post}}$ and underestimated if $\overline{T_{pre}} < \overline{T_{post}}$, due to increasing forward and backward drift, respectively, of the strobed phases within successive post-return cycles.

Phase shift, $\Delta\phi$, the amount that a given trial's post-return rhythm was shifted relative to its pre-perturbation rhythm, was then defined as:

$$\Delta\phi = \left(\overline{\phi_{new,corrected}} - \overline{\phi_{old}} + 0.5\right) \text{(modulo 1)} - 0.5 \quad (2)$$

Thus, $\Delta\phi > 0$ denotes phase advance, $\Delta\phi < 0$ denotes phase delay, and $\Delta\phi = 0$ denotes no phase shift. The modular arithmetic puts these values on the interval [-0.5,+0.5].

The *relative phase* of the bilabial and laryngeal trajectories was defined operationally by using successive peak laryngeal openings as "strobe" events in each corresponding bilabial cycle. Relative phase for each laryngeally-strobed bilabial cycle could then be viewed as:

$$\phi_{rel} = \frac{\left[\begin{array}{c}\text{(time of } i^{th} \text{ laryngeal event)} \\ -\text{(time of the preceding bilabial peak)}\end{array}\right]}{\text{(period of the strobed bilabial cycle)}} \quad (3)$$

Computationally, average relative phase for each perturbed trial's strobed bilabial pre-perturbation cycles was defined as $\overline{\phi_{rel,\,pre}}$ = [bilabial $\overline{\phi_{old}}$ – laryngeal $\overline{\phi_{old}}$] (modulo 1); average relative phase for the strobed bilabial post-return cycles was defined as $\overline{\phi_{rel,\,post}}$ = [bilabial $\overline{\phi_{new,\,corrected}}$ – laryngeal $\overline{\phi_{new,\,corrected}}$] (modulo 1). Steady-state shifts in relative phase for each trial, $\Delta\phi_{rel}$, were

taken as the post-/pre-perturbation difference in relative phase ($\phi_{rel,\ post} - \phi_{rel,\ pre}$), with the same modular arithmetic applied as above (i.e., of the same form as Eq. 2). Thus, a positive (or negative) shift in relative phase indicates a perturbation-induced delay (or advance) of peak laryngeal opening with respect to the bilabial cycle.

The same cycle types and experimental measures were obtained for the control (no perturbation) trials, where calculations were anchored to the end of a randomly timed, but not delivered, perturbation (hereafter referred to as a "phantom perturbation"). For the perturbed trials, the measures of phase shift and shifts in relative phase were converted to (experimental–control) difference scores, where control = the session-specific control values computed for each measure averaged across all of the phantom-perturbation times. In order to examine the system's sensitivity to perturbations delivered in different portions of the cycle, these difference scores were partitioned into five bins and averaged according to a normalized measure of perturbation-delivery time defined by (perton–t_{on1})/prepert, where *perton* = onset time of perturbation, t_{on1} = onset time of the first-perturbation bilabial cycle, and prepert = the average duration of the bilabial pre-perturbation cycles for the trial. Note that since the first-perturbation bilabial cycle could be longer than the mean pre-perturbation cycle duration, the final time bin extends from 0.88 to 1.10.[1]

For each difference measure, separate sets of *t*-tests were computed for each perturbation bin to test whether the measures differed from zero. To protect against an elevated Type-I error rate due to multiple comparisons across perturbation bins, α - levels were selected by dividing 0.05 by the number of bins. All reported significant, protected *t*-tests were significant at $P<0.05$ or better. In order to detect differences across the time bins in all of our dependent measures, one-way repeated-measures ANOVAs with "bin" as a factor were performed. Post-hoc Tukey tests based on these ANOVAs were computed, and all reported significances were significant at $P<0.05$ or better.

Repetitive sequences: transient analyses

In the transient analyses of the repetitive sequences, the timing/phasing changes that occurred during the first and second perturbation cycles (see Fig. 3) were analyzed for both bilabial and laryngeal trajectories. Additionally, transient behavior of the relative phasing between bilabial and laryngeal trajectories was also examined. As with the steady-state analyses described earlier, control measures were calculated for the first and second perturbation cycles that were defined by randomly timed, but not delivered, "phantom" perturbations.[2]

For each trial, the duration changes of the first-perturbation cycles (dur_1) for bilabial and laryngeal data were expressed separately as normalized difference scores with respect to the trial's average pre-perturbation cycle period for the respective trajectory types using the formula (dur_1–prepert)/prepert. Similarly, the duration changes of the second-perturbation cycles (dur_2) for bilabial and laryngeal data were expressed separately for each trial as normalized difference scores with respect to the trial's average pre-perturbation cycle period using the formula (dur_2–prepert)/prepert. Transient relative phases ($\phi_{rel,pert-i}$, i=1, 2) were defined by strobing the first and second bilabial perturbation cycles at peak laryngeal opening using the formula: (time of peak laryngeal opening from the onset of the i^{th} bilabial perturbation cycle)/(period of the strobed i^{th} bilabial perturbation cycle). Finally, transient shifts in relative phase for each trial ($\Delta\phi_{rel,pert-i}$, i=1, 2) were taken as the differences between the transient relative phases and the mean pre-perturbation relative phase ($\Delta\phi_{rel,pert-i} - \phi_{rel,\ pre}$), again submitted to modular arithmetic using the same form as Eq. 2.

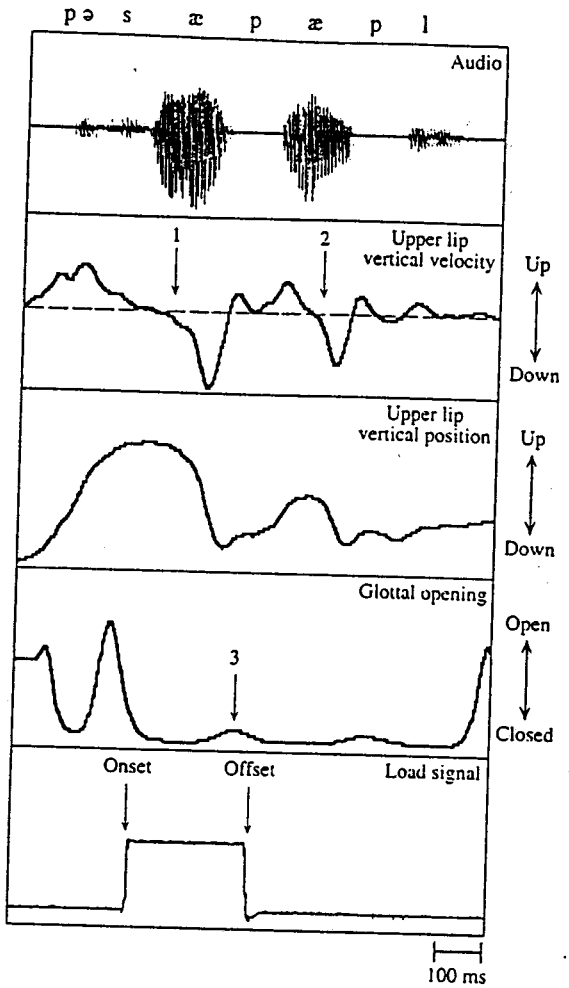

Fig. 4 Plot of audio signal, upper lip, laryngeal opening, and load signal during a representative perturbed discrete trial. *Events 1* and *2* denote onsets of the upper-lip lowering movements for the second and third /p/ closures, respectively; *event 3* denotes the onset of glottal adduction associated with the second /p/

The articulator-specific normalized duration changes, as well as shifts in relative phase, were converted to (experimental–control) difference scores, where control = the session-specific control values computed for each measure averaged across all of the "phantom perturbation" times. These difference scores were partitioned into five bins and averaged according to normalized measures of perturbation delivery time defined by (perton–t_{on1})/prepert, where *perton* = onset time of perturbation; t_{on1} = onset time of the bilabial (for the bilabial duration changes, and transient shifts in relative phase) or laryngeal (for the laryngeal duration changes) first-perturbation cycle; and prepert = the average duration of the bilabial (for the bilabial duration changes, and transient shifts in relative phase) or laryngeal (for the laryngeal duration changes) pre-perturbation cycles for the trial. Protected *t*-tests were computed for

[1] Out of 300 total repetitive trials, for subject ES 180 perturbed trials and 38 control trials were included in the steady-state analyses; 30 trials were lost due to equipment malfunctioning during session 1, and the remaining excluded trials were due to noisy, unanalyzable pre-perturbation or post-return portions of the bilabial or laryngeal trajectories (or both) that did not allow reliable peak detection. For subject AL, 221 perturbed trials and 60 control trials were included in the analyses; all excluded trials were due to noisy, unanalyzable signals.

[2] For subject ES, 12 perturbed repetitive trials in addition to those excluded from the steady-state analyses (see Footnote 1) were excluded from the transient analyses, due to unanalyzable transient portions of the bilabial or laryngeal trajectories (or both). For subject AL, an additional 8 perturbed repetitive trials were excluded. No additional control trials were excluded from the transient analyses for either subject.

each perturbation bin in each data set to test whether the perturbation-induced changes differed from zero.

Discrete sequences

In the discrete sequences, each trial consisted of the utterance /pə'sæpæpl/. Oral articulatory intervals were identified in the upper-lip velocity signal. The upper-lip signal was used since it is mechanically relatively unaffected by the load applied to the lower lip. To obtain a criterion for movement onsets, the maximum peak velocities of upper-lip raising and lowering recorded during the experimental session were identified. Onsets were then identified algorithmically as the point at which the upper-lip velocity reached 10% of the maximum peak velocity. Peak glottal openings were identified using a simple peak-picking procedure. The audio signal and the articulatory and loading events for a single representative trial are shown in Fig. 4[3]. Two temporal articulatory intervals were measured. The first interval (*syllable duration*) was defined between the onsets of the two upper-lip lowering movements for the second and third /p/ closures (events 1 and 2 in Fig. 4). Normalized perturbation-induced changes in syllable duration were computed according to (syllable duration–$\overline{\text{control}_d}$), where $\overline{\text{control}_d}$ = the mean value of syllable duration for the control (non-perturbed) discrete trials of the respective session. The second temporal interval measured was the interval between the onset of upper-lip lowering for the second /p/ closure (event 1 in Fig. 4) and the onset of glottal adduction (event 3 in Fig. 4). This measure was used to describe the *relative phase* between the oral and laryngeal events by dividing it by the syllable duration interval. Perturbation-induced shifts in relative phase were computed according to (relative phase–$\overline{\text{control}_d}$), where $\overline{\text{control}_d}$ = the mean control, non-perturbed value of relative phase for the discrete trials of the respective session.

The articulatory measures (changes in syllable duration and relative phase) were partitioned into four bins and averaged according to a normalized measure of perturbation delivery time. For historical reasons (e.g., Gracco and Abbs 1989), this normalized time base was defined by the formula (*pertoff* – t_{offd})/$\overline{\text{control}_d}$, where *pertoff* = offset time of perturbation (load offset in Fig. 4), t_{offd} = offset time of discrete syllable duration interval (event 2 in Fig. 4), and $\overline{\text{control}_d}$ = average duration of discrete syllable duration intervals from the unperturbed control trials for the respective session. Protected *t*-tests were computed for each articulatory interval, comparing whether the mean normalized syllable duration changes and shifts in relative phase in each of the four perturbation bins differed from zero.

Comparison of discrete and repetitive sequences

In the present study and in previous work (e.g., Gracco and Abbs 1989), step-pulse perturbations applied during discrete speech sequences were one syllable long and were experimentally timed so that perturbation offset occurred at various times within the syllable of interest (defined between maximum vocalic lip openings). In our repetitive task, step-pulse perturbations were also approximately one cycle long (again defined between maximum lip openings). This means that perturbation onset occurred, by definition, during the first-perturbation cycle and offset generally occurred during the following, second-perturbation cycle. Consequently, the two articulatory measures studied in the repetitive sequences most comparable to those studied in the discrete sequences are: (1) the duration of the second-perturbation lip-aperture cycle (dur_2), analogous to the syllable-duration interval in the discrete sequences; and (2) the lip-larynx relative phase for the second-perturbation cycle, analogous to the relative phase of the laryngeal peak inside

[3] Of 300 total discrete trials, for subject ES 211 perturbed trials and 66 control trials were included in the analyses; the remaining trials were excluded due to noisy, unanalyzable portions of the bilabial or laryngeal trajectories (or both). For subject AL, 236 perturbed trials and 63 control trials were included in the analyses.

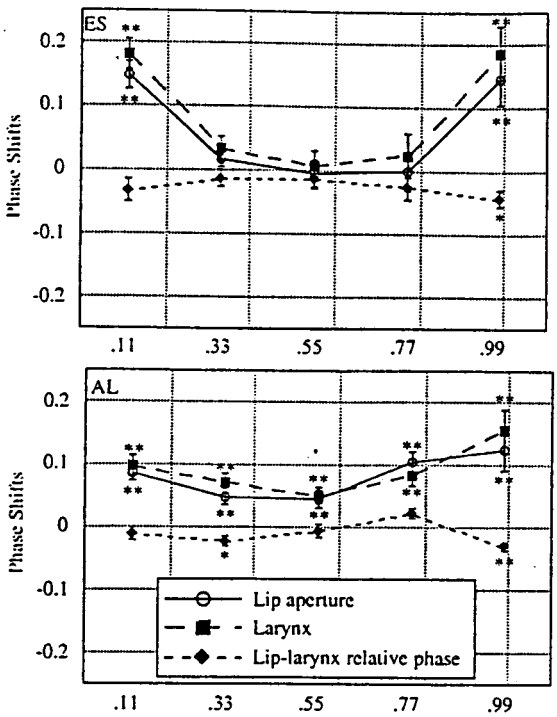

Fig. 5 Difference scores ([experimental–$\overline{\text{control}}$]/$\overline{\text{control}}$) for steady-state phase shifts for the lip-aperture (*open circles*) and laryngeal (*filled squares*) trajectories as well as shifts in relative phase (*filled diamonds*) between lip and laryngeal trajectories, binned and averaged according to normalized perturbation onset time [(perton–t_{on1})/prepert], where t_{on1} denotes onset time of the first perturbation lip-aperture cycle and prepert denotes the average duration of the pre-perturbation lip-aperture cycles. *Bin labels* represent the centers of these bins. *Error bars* denote standard errors. *Top panel* ES's data, *bottom panel* AL's data. * and ** denote, respectively, $P<0.05$ and $P<0.01$ significance levels for the protected *t*-tests

the syllable duration interval in discrete sequences. In order to compare the effects of perturbation on these repetitive measures with their discrete counterparts, two measures were computed for the second-perturbation cycles. The first reflected perturbation-induced changes in second-perturbation cycle durations and was defined by (dur_2–$\overline{\text{control}_r}$)/$\overline{\text{control}_r}$, where $\overline{\text{control}_r}$ = the average duration of the "second-perturbation cycles" (defined by randomly timed, but not delivered, "phantom" perturbations) for the control, non-perturbed trials of the respective session. The second measure reflected perturbation-induced shifts in lip-larynx relative phase and was defined by (relative phase–$\overline{\text{control}_r}$), where $\overline{\text{control}_r}$ = the mean control, non-perturbed value of relative phase for the "second-perturbation cycles" of the respective session.

These normalized repetitive difference measures were binned and averaged within a set of four time bins comparable to those used to partition the discrete data. Thus, the bins were defined according to a normalized measure of perturbation delivery time, corresponding to that used in the discrete sequences: (*pertoff* – t_{offr})/$\overline{\text{control}_r}$, where *pertoff* is the perturbation offset time, t_{offr} is the offset time of the second-perturbation cycle, and $\overline{\text{control}_r}$ is the average duration of unperturbed control "second-perturbation cycles" for the respective session. Protected *t*-tests were computed for the repetitive data, comparing whether the mean normalized duration changes and shifts in relative phase in each of the four perturbation bins differed from zero.

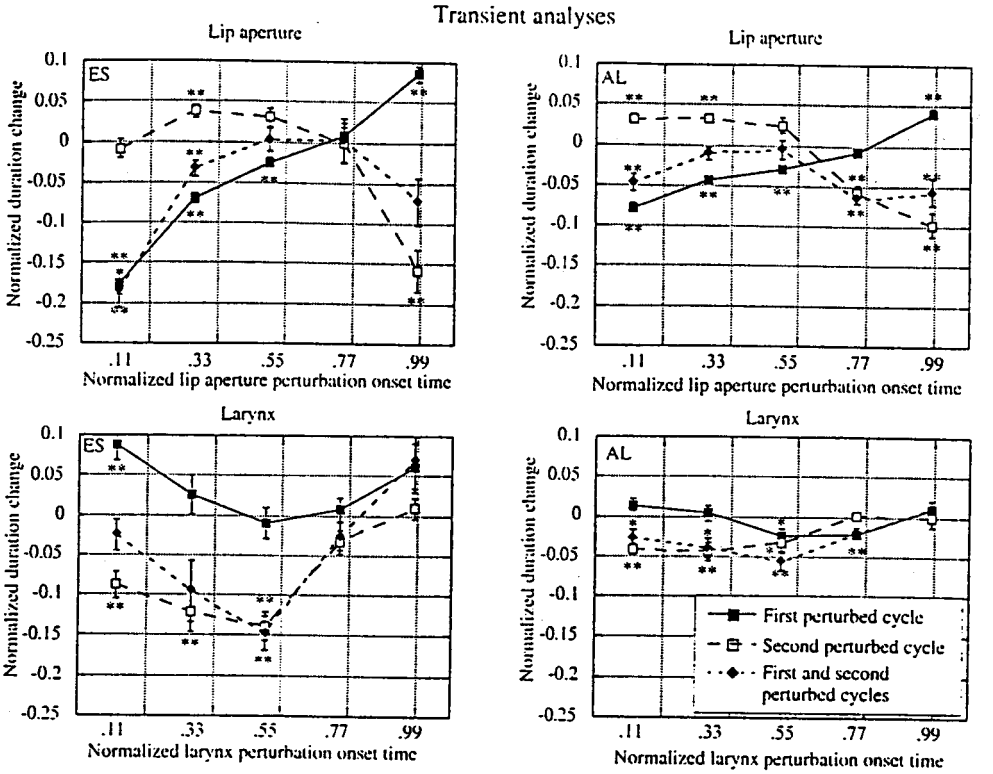

Fig. 6 Difference scores ([experimental–control]/control) for fractional duration changes of the lip aperture (*top row*) and laryngeal (*bottom row*) trajectories' first perturbed cycles (*filled squares*), second perturbed cycles (*open squares*), and first-plus-second perturbed cycles (*filled diamonds*), binned and averaged according to normalized perturbation onset time [(perton–t_{on1})/prepert], where t_{on1} and prepert denote, respectively, the onset time of the first perturbation cycle and the average duration of the pre-perturbation cycles of the analyzed trajectory (lip aperture or larynx). *Bin labels* represent the centers of these bins. *Error bars* denote standard errors. *Right column* ES's data, *left column* AL's data. * and ** as in Fig. 5

Results

Repetitive sequences: phase-resetting (steady-state) analyses

Figure 5 shows the results of the steady-state analyses of bilabial and laryngeal phase shifts as well as shifts in relative phases for subjects ES (top panel) and AL (bottom panel). For ES, protected *t*-tests showed that the bilabial and laryngeal rhythms were only significantly phase advanced relative to the no-perturbation control trials in the 0.11 and 0.99 time bins. The bilabial pattern replicates the phase-resetting results found in earlier studies of this speaker, which focused only on bilabial behavior (Saltzman 1992; Saltzman et al. 1991). Protected *t*-tests also only showed a significant shift (negative) in the relative phasing of lips and larynx relative to control values in the 0.99 time bin. One-way ANOVAs performed separately for the bilabial, laryngeal, and relative phase data revealed main effects for the bilabial [$F(4,175)=7.9$, $P<0.001$] and laryngeal [$F(4,175)=8.62$, $P<0.001$] phase shifts. Post-hoc Tukey tests indicated that, for this subject's bilabial and laryngeal data, phase shifts were greater in bins 0.11 and 0.99 than in bins 0.33, 0.55, and 0.77. There was no significant main effect for shifts in relative phase [$F(4.175)=0.54$, $P=0.7$].

These results indicate that phase-resetting of the bilabial and laryngeal trajectories occurred for ES, but was temporally localized to a "sensitive period" immediately preceding (time bin 0.99) and following (time bin 0.11) peak bilabial opening. During this time, the downwardly directed lower-lip perturbation was delivered near the initiation of the actively controlled bilabial closing gesture for /p/ and acted, thereby, to oppose that gesture. That is, resetting occurred (roughly) during the acceleration portion of the closing gesture, i.e., during the intervals of opening deceleration and closing acceleration surrounding the peak bilabial opening (Kawato, personal communication). Additionally, a significant shift in lip-larynx relative phasing was observed (time bin 0.99), although this shift was an order of magnitude smaller than the individual phase shifts of the bilabial and laryngeal gestures.

For subject AL's bilabial and laryngeal data, protected *t*-tests showed significant phase advances in all time bins. Significant (negative) shifts in the relative phasing of lips and larynx occurred only in the .33 and .99 time bins. As for subject ES, one-way ANOVAs performed separately for the bilabial, laryngeal, and relative phase data revealed significant main effects for the bilabial ($F[4,216]=3.77$, $P<.01$) and laryngeal ($F[4,216]=3.5$, $P<.01$) phase shifts. Post-hoc Tukey tests indicated that, for this subject's bilabial and laryngeal data, phase shifting was greater in bin .99 than in bins .33 and .55. There was also a significant main effect for shifts in relative phase ($F[4,216]=5.15$, $P<.001$); post-hoc Tukey tests indicated that these shifts were greater in bin .77 than in bins .33 and .99.

These results indicate that, similar to ES, phase resetting occurred for AL, which was maximal near the time of peak bilabial opening, and that shifts in lip-larynx relative phasing occurred that were an order of magnitude smaller than the individual phase shifts of the bilabial and laryngeal gestures. Unlike ES, however, AL showed: (1) a significant level of bilabial and laryngeal resetting throughout the syllable cycle; (2) maximal sensitivity in such resetting

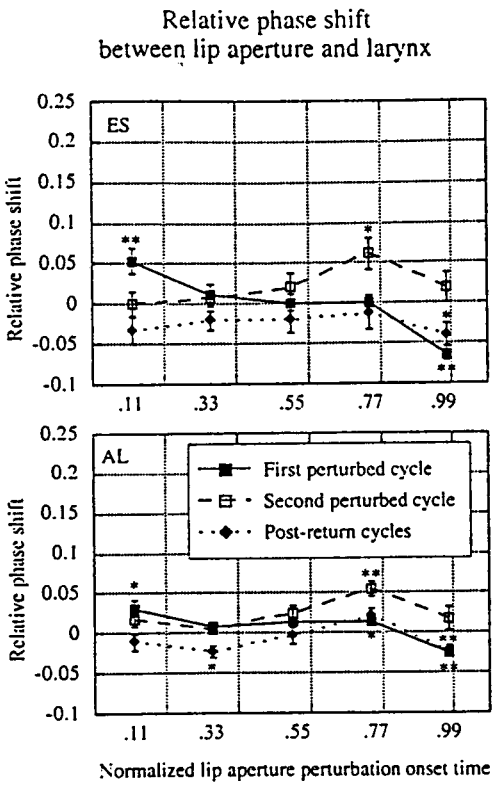

Fig. 7 Difference scores ([experimental–$\overline{\text{control}}$]/$\overline{\text{control}}$) for shifts in lip-larynx relative phase for first perturbed cycles (*filled squares*), second perturbed cycles (*open squares*), and post-return cycles (*filled diamonds*), binned and averaged according to normalized lip aperture perturbation onset time [(perton–t_{on1})/$\overline{\text{prepert}}$]. *Bin labels* represent bin centers. *Error bars* denote standard errors. *Top panel* ES's data, *bottom panel* AL's data. * and ** as in Fig. 5

only when perturbations were delivered immediately prior to peak opening (time bin 0.99 only); and (3) shifts in lip-larynx relative phase in time bin 0.33 as well as bin 0.99.

Repetitive sequences: transient analyses and the relation between steady-state and transient analyses

Figure 6 shows the bilabial and laryngeal normalized cycle duration changes for the first- (solid squares) and second- (open squares) perturbation cycles for subjects ES (right column) and AL (left column). Figure 7 shows shifts in relative phase for the first- (solid squares) and second- (open squares) perturbation cycles for subjects ES (top panel) and AL (bottom panel). The results of protected *t*-tests assessing the significance of the perturbation-induced changes relative to zero are also displayed.

For subject ES, one-way ANOVAs performed on the first-perturbation cycles revealed significant main effects for the bilabial [$F(4,169)=90.58$, $P<0.001$] and laryngeal [$F(4,169)=3.77$, $P<0.01$] cycle-duration changes and for the relative phase shifts [$F(4,163)=11.24$, $P<0.001$]; for the second-perturbation cycles, there were significant main effects for the bilabial [$F(4,164)=18.85$, $P<0.001$] and laryngeal [$F(4,164)=9.62$, $P<0.001$] duration changes, but not for the relative phase shifts [$F(4,163)=2.2$, $P=0.071$]. Post-hoc Tukey tests indicated that: (1) for the bilabial first-perturbation cycles, the only time bins whose duration changes did not significantly differ from each other were bins 0.55 and 0.77. All other bins differed from one another; (2) for the bilabial second-perturbation cycles, the duration changes for bin 0.99 were sig-

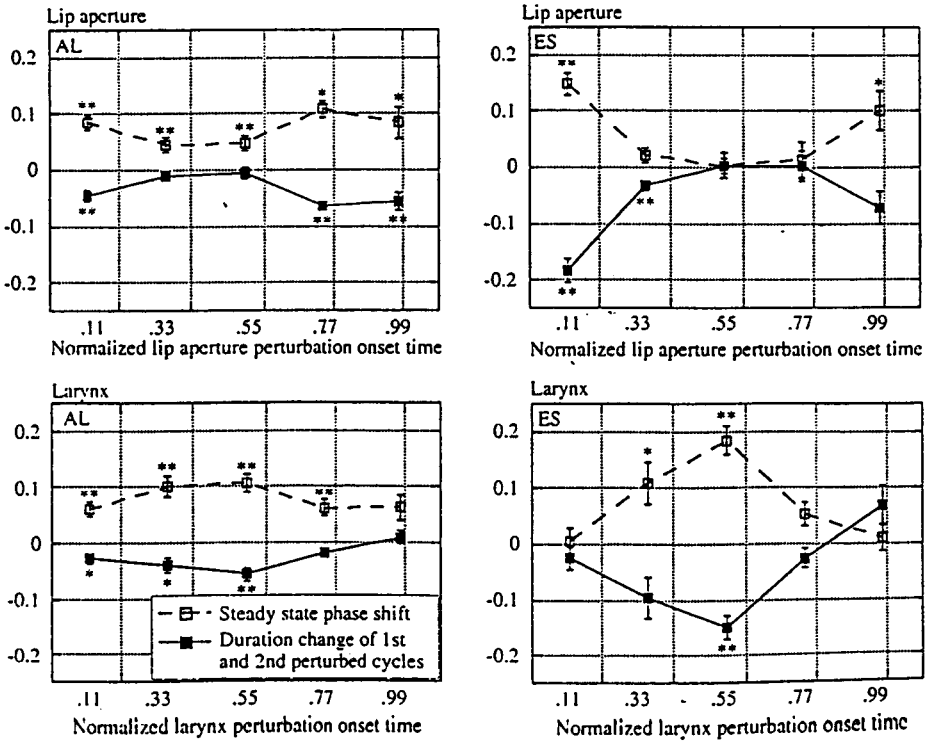

Fig. 8 Difference scores ([experimental–$\overline{\text{control}}$]/$\overline{\text{control}}$) for lip-aperture (*top row*) and laryngeal (*bottom row*) phase shifts (*open squares*) and first-plus-second perturbed cycles' fractional duration changes (*filled squares*), binned and averaged according to normalized perturbation onset time [(perton–t_{on1})/$\overline{\text{prepert}}$], where t_{on1} and $\overline{\text{prepert}}$ denote, respectively, the onset time of the first perturbation cycle and the average duration of the preperturbation cycles of the analyzed trajectory (lip aperture or larynx). *Bin labels* represent the centers of these bins. *Error bars* denote standard errors. *Right column* ES's data, *left column* AL's data. * and ** as in Fig. 5

nificantly different from all other bins; (3) for the laryngeal first-perturbation cycles, the duration changes for bin 0.11 were significantly different from bins 0.55 and 0.77; (4) for the laryngeal second-perturbation cycles, the duration changes for bin 0.99 were significantly different from bins 0.11, 0.33, and 0.55. The duration changes for bin 0.77 were significantly different from bins 0.33 and 0.55; and (5) for the first-perturbation cycles, relative phase shifts were more negative in bin 0.99 than in all other bins. Additionally, the relative phase shifts were more positive in bin 0.11 than in bins 0.55 and 0.77.

For subject AL, one-way ANOVAs performed on the first-perturbation cycles revealed significant main effects for the bilabial [$F(4,208)=66.53$, $P<0.001$] and laryngeal [$F(4,208)=4.98$, $P<0.001$] duration changes and for the relative phase shifts [$F(4,208)=7.96$, $P<0.001$]; for the second-perturbation cycles, there were significant main effects for the bilabial [$F(4,208)=44.32$, $P<0.001$] and laryngeal [$F(4,208)=5.67$, $P<0.001$] duration changes as well as for the relative phase shifts [$F(4,204)=4.88$, $P<0.001$]. Post-hoc Tukey tests indicated that: (1) for the bilabial first-perturbation cycles, the only time bins whose duration changes did not significantly differ from each other were bins 0.33 and 0.55; all other bins differed from one another; (2) for the bilabial second-perturbation cycles, the duration changes for bins 0.77 and 0.99 were significantly different from all other bins and from each other ; (3) for the laryngeal first-perturbation cycles, the duration changes for bin 0.11 were significantly different from bins 0.55 and 0.77; (4) for the laryngeal second-perturbation cycles, the duration changes for bin 0.77 were significantly different from bins 0.11, 0.33, and 0.55; (5) for the first-perturbation cycles, relative phase shifts were more negative in bin 0.99 than in all other bins; additionally, the relative phase shifts were more positive in bin 0.11 than in bin 0.33; and (6) for the second-perturbation cycles, relative phase shifts were more positive in bin 0.77 than in bins 0.33 and 0.11.

We examined the hypothesis that most of the bilabial and laryngeal phase shifts observed in the steady-state were attributable to duration changes induced during the application of the perturbation. To do this, we first summed the normalized duration changes for each trial's first- and second-perturbed cycles, and then binned and averaged them using the same normalized time bases as in our previous analyses for these cycles [see Fig. 6 (filled diamonds) and Fig. 8 (filled squares)]. The results of protected *t*-tests assessing the significance of the perturbation-induced changes relative to zero are also displayed.

For subject ES, one-way ANOVAs revealed significant main effects for the bilabial [$F(4.164)=12.86$, $P<0.001$] and laryngeal [$F(4.164)=8.45$, $P<0.001$] duration changes. Post-hoc Tukey tests indicated that: (1) for the bilabial summed cycles, time bin 0.11 differed significantly from all other bins; and (2) for the laryngeal summed cycles, the duration changes for bin 0.33 were significantly different from bin 0.99, and those for bin 0.55 differed significantly from bins 0.11, 0.77, and 0.99.

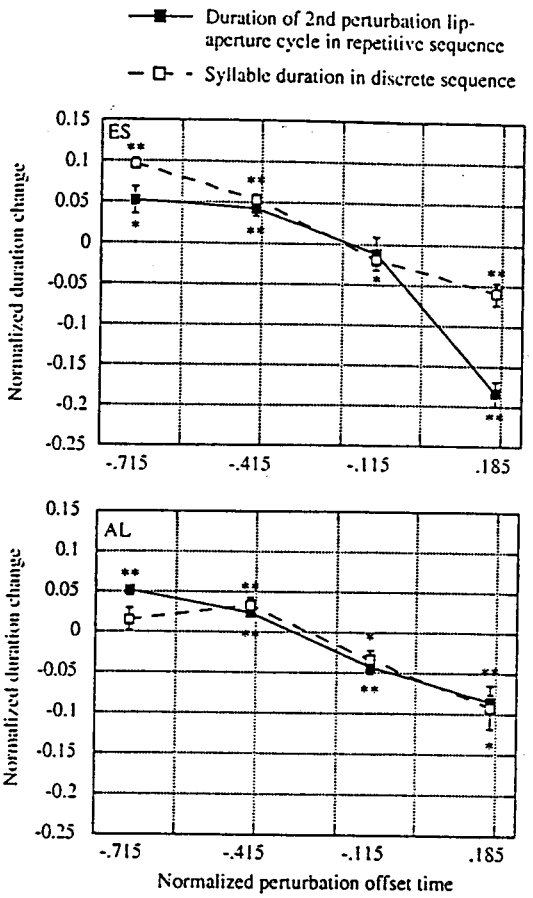

Fig. 9 Normalized duration changes of lip-aperture syllable (discrete data, *open squares*, defined by [{experimental−$\overline{\text{control}_d}$}/$\overline{\text{control}_d}$]) and second perturbed cycle (repetitive data, *filled squares*, defined by [{experimental−$\overline{\text{control}_r}$}/$\overline{\text{control}_r}$]), binned and averaged according to their respective normalized time bases ([{pertoff−t_{offd}}/$\overline{\text{control}_d}$] for discrete data, [{pertoff−t_{offr}}/$\overline{\text{control}_r}$] for repetitive data). *Bin labels* represent bin centers. *Error bars* denote standard errors. *Top panel* ES's data, *bottom panel* AL's data. * and ** as in Fig. 5

For subject AL, one-way ANOVAs revealed significant main effects for the bilabial [$F(4.208)=6.87$, $P<0.001$) and laryngeal [$F(4.208)=2.97$, $P<0.05$) duration changes. Post-hoc Tukey tests indicated that: (1) for the bilabial summed cycles, time bins 0.33 and 0.55 differed significantly from bins 0.77 and 0.99; and (2) for the laryngeal summed cycles, the duration changes for bin 0.55 were significantly different from bin 0.99.

We then compared the steady-state data curves (Fig. 8, open squares)[4] with the summed perturbation cycle data curves (Fig. 8, filled squares) for the bilabial and la-

[4] In order to compare transient and steady-state behaviors in the repetitive trials, the present steady-state data analyses were performed using the same time bases as in Fig. 6. Additionally, these analyses were performed on the same sets of trials for each subject as those used in the transient analyses (see Footnote 2). Thus, there are slightly fewer trials used here in the steady-state analyses than were included in the steady-state analyses described in the prior data-analysis procedures section entitled "Repetitive sequences: phase-resetting (steady-state) analyses".

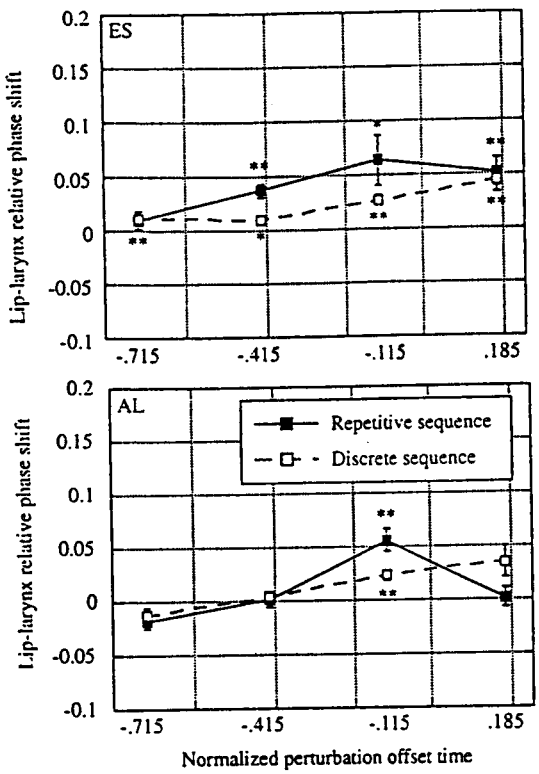

Fig. 10 Shifts in lip-larynx relative phase for lip-aperture syllable (discrete data, *open squares*, defined by [relative phase–$\overline{control_d}$]) and second perturbed cycle (repetitive data, *filled squares*, defined by [relative phase–$\overline{control_r}$]), binned and averaged according to their respective normalized time bases ([{pertoff—t_{offd}}/$\overline{control_d}$] for discrete data, [{pertoff – t_{offr}}/$\overline{control_r}$] for repetitive data). *Bin labels* represent bin centers. *Error bars* denote standard errors. *Top panel* ES's data, *bottom panel* AL's data. * and ** as in Fig. 5

ryngeal data. Note that these curves are virtually mirror images of each other. Given the opposite sign conventions used to define the steady-state phase shifts and summed cycle-duration changes, these curves are in essence identical, indicating that the steady-state phase shifts are indeed induced primarily in the first two perturbation cycles.

Discrete sequences: comparison with repetitive data

Changes in syllable duration (Fig. 9, open squares) and relative phases (Fig. 10, open squares) were analyzed for the discrete utterances. Additionally, we analyzed the two articulatory measures from those repetitive sequences (duration change of the second-perturbation cycle: Fig. 9, filled squares; change in lip-larynx relative phase for the second-perturbation cycle: Fig. 10, filled squares) that are most comparable to the above discrete measures (see the earlier data-analysis procedures section entitled "Comparison of discrete and repetitive sequences"). Inspection of the data curves and of their patterns of significance indicates that: (1) both subjects displayed similar durational lengthening and shortening in the discrete and repetitive utterances (Fig. 9); and (2) patterns of changes in relative phase for the discrete and repetitive utterances were similar within subjects, but possibly different across subjects (Fig. 10). These results suggest that the behaviors observed in the (relatively unnatural) rhythmic sequences and the (more natural) discrete sequences are indeed governed by a common underlying dynamic organization.

Discussion

Our findings led us to several conclusions. First, the steady-state analyses of both speakers' repetitive utterances indicated that "permanent" phase shifts existed for both the lips and the larynx after the system returned to its pre-perturbation steady-state rhythm. These results support the hypothesis that central intergestural dynamics can be reset by peripheral articulatory events. Such resetting was strongest when the downwardly directed lower-lip perturbation was delivered near the initiation or acceleration portion of the actively controlled bilabial closing gesture for /p/. Second, analyses of the transient portions of the perturbed cycles of the repetitive utterances indicated that the perturbation-induced steady-state phase shifts were almost totally attributable to changes occurring during the first two perturbed cycles. Third, in addition to the steady-state shifts in the timing between successive bilabial closing and laryngeal devoicing gestures for /p/, steady-state shifts in the *relative* phasing of these gestures were also demonstrated. However, the individual temporal shifts of the bilabial and laryngeal gestures were an order of magnitude larger than the relative temporal shift between these gestures, and the lips and larynx appeared to be phase-advanced as a relatively coherent unit. Thus, these results not only demonstrate a resetting of a central "clock" for these utterances, but also imply that intergestural temporal cohesion is greater within phonemes (i.e., between labial and laryngeal gestures during each /p/) than between phonemes (i.e., between labial or laryngeal gestures in the /p/s of successive syllables), as has been hypothesized by Byrd (1996), Löfqvist (1991), Nittrouer, et al. (1988), and Saltzman and Munhall (1989) (see also Gracco and Löfqvist 1994). Finally, the transient changes in speech timing induced by perturbations in the discrete sequences are similar to those that occur in comparable portions of the repetitive sequences, and, thus, both the (more natural) discrete utterances and the (less natural) repetitive utterances appeared to share a common dynamic basis.

What sort of dynamical system could account for these temporal patterns? One candidate, which we have described previously (e.g., Byrd et al. 1998, in press; Fowler and Saltzman 1993; McGowan and Saltzman 1995; Saltzman 1991, 1995; Saltzman and Munhall 1989; Saltzman et al. submitted), is a two-component model in which: (1) *interarticulator* dynamics account for the coordination among component articulators (e.g., upper and

lower lip) during single and co-produced speech gestures; and (2) *intergestural* dynamics account for the temporal patterning of activity among the gestures that compose a given utterance. Our steady-state phase-resetting data support the hypothesis that such intergestural dynamics provide a "clocking" mechanism that can be reset (phase-advanced) by peripheral articulatory events, and that such resetting is strongest when gesture-opposing perturbations are delivered near the gesture's initiation. Thus, these data imply that intergestural timing networks for speech do not unidirectionally drive the articulatory periphery. Rather, intergestural and interarticulatory dynamics must be coupled bidirectionally, so that feedback information can influence the intergestural clock in a manner that is sensitive to articulatory state (e.g., Kay et al. 1991). In particular, these data provide strong evidence that gestural activation patterns are not rigidly specified over a given sequence, but evolve fluidly and flexibly over the course of an ongoing sequence as implicit consequences of the dynamics of the entire multilevel system.

In our previous simulations of temporal patterning in speech (Saltzman 1995; Saltzman et al. 1998, in press), which we have not yet extended to a modeling of the present phase-resetting data, we used a recurrent connectionist network (i.e., Jordan's "sequential network" architecture, described in Jordan 1986, 1988, 1990, 1992; and in Jordan and Rumelhart 1992) in order to define an intergestural dynamic system that could serve as a programmable, utterance-specific central pattern generator or clock. In our networks, *output* units define patterns of gestural activation, and the activity of *state* units provide a time scale that is intrinsic to the intended sequence. Additionally, each output unit's activity is fed back to a single state unit that acts as a linear first-order filter. In related models (e.g., Bailly et al. 1991; Jordan 1988; Laboissière et al. 1991), pairs of state units define linear second-order filters, which act as internal oscillators and provide a clock-like representation of network time.

Our experimental results as well as those on phase transitions in speech production (Kelso et al. 1986a, b; Tuller and Kelso 1991) are consistent with the use of state unit oscillators in models of intergestural dynamics. Additionally, however, these findings suggest that separate oscillators are associated with each gestural unit, and that these oscillators are mutually coupled, nonlinear limit cycles. Thus, in our data, the observed pattern of intergestural activity during unperturbed repetitive trials would correspond to an associated pattern of utterance-specific synchronization (entrainment) and relative phasing among such state-oscillators. In perturbed cases, these oscillators would display steady-state phase-resetting. They would also display steady-state shifts of relative phasing if either: (1) the system observation time (approximately 20 syllables) was shorter than the relaxation time required to return to the system's pre-perturbation entrainment pattern; or (2) the initially observed relative phasing was simply one value in a phase window (Byrd 1996) or interval of allowable relative phases.

The second type of data consistent with a model incorporating entrained nonlinear state-oscillators is provided by rate-scaling experiments of intergestural phase transitions (Kelso et al. 1986a, b; Tuller and Kelso 1991). In these experiments, continuous increases in speaking rate produced discontinuous transitions of intergestural phasing. For example, 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 a modal pattern, which becomes unstable with increasing rates, to another modal pattern that retains its stability (e.g., Haken et al. 1985). The implications for model development are that state-unit oscillators should be of the type that has been shown to exhibit the above patterns of synchronization and bifurcation. Such systems of coupled, opponent-pair limit cycles have been described (Cohen et al. 1992; Grossberg et al. 1997; Nagashino and Kelso 1991, 1992; Pribe et al. 1997) and are logical candidates for defining the set of state-unit oscillator pairs in dynamic models of these behaviors.

Acknowledgements This work was supported by NIH Grants DC-00121, DC-00865, and DC-03663 from the National Institute of Deafness and Other Communication Disorders, by Grant BNS-88–20099 from the National Science Foundation, and by Esprit-BR Project 6975-Speech Maps through Nutek Grant P55–1. The authors wish to thank Frank Guenther and an anonymous reviewer for their helpful reviews of the original manuscript; in particular, Eq. 2 was provided by the latter reviewer as an elegant substitute for an originally unwieldy set of equivalent expressions.

Appendix: derivation of Eq. 1

Assume that the pre-perturbation and post-return cycles have constant periods equal to T_{pre} and T_{post}, respectively, and recall that the series of new phases, $\phi_{i\text{-new}}$ (i=1, 2, ..., n), is defined by strobing successive post-return cycles using a strobe interval equal to T_{pre}. If T_{post}=T_{pre}, the successive new phase values will be identical. However, if $T_{post} \neq T_{pre}$, these values will drift according to:

$$\phi_{i-new} = \left(\frac{\Delta t_1 + (i-1)(T_{pre} - T_{post})}{T_{post}} \right) (\text{modulo } 1) \qquad \text{(A1)}$$

where Δt_1 is the strobe time for the first post-return cycle measured from the beginning of that cycle, and $\Delta t_1 / T_{post}$ is the corresponding new phase value, ϕ_{1-new}. Thus, the average new phase, $\overline{\phi_{new}}$, is defined as:

$$\overline{\phi_{new}} = \left(\frac{1}{n} \sum_{i=1}^{n} \phi_{i-new} \right) (\text{modulo } 1) \qquad \text{(A2)}$$

Substituting from Eq. A1, we get:

$$\overline{\phi_{new}} = \left(\phi_{1-new} - \left(\frac{T_{pre} - T_{post}}{T_{post}}\right) + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{i\left(T_{pre} - T_{post}\right)}{T_{post}}\right)\right) \quad \text{(modulo 1)} \qquad \text{(A3)}$$

Using the summation properties of arithmetic series, this simplifies to:

$$\overline{\phi_{new}} = \left(\phi_{1-new} + \left(\frac{n-1}{2}\right)\left(\frac{T_{pre} - T_{post}}{T_{post}}\right)\right)(\text{modulo 1}) \qquad \text{(A4)}$$

Finally, substituting average values, $\overline{T_{post}}$ and $\overline{T_{pre}}$, for the assumed constant values T_{post} and T_{pre}, respectively, and defining $\overline{\phi_{new;\ corrected}} = \phi_{1-new}$, we arrive at Eq. 1 in the text.

References

Abramson AS (1977) Laryngeal timing in consonant distinctions. Phonetica 34:295–303

Baer T, Löfqvist A, McGarr NS (1983) Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. J Acoust Soc Am 73:1304–1308

Bailly G, Laboissière R, Schwartz JL (1991) Formant trajectories as audible gestures: an alternative for speech synthesis. J Phonetics 19:9–23

Byrd D (1996) A phase window framework for articulatory timing. Phonology 13:139–169

Byrd D, Narayanan S, Kaun A, Saltzman E (1998) Phrasal signatures in articulation. In: Broe M, Pierrehumbert J (eds) Papers in laboratory phonology V (in press)

Cohen MA, Grossberg S, Pribe C (1992) A neural pattern generator that exhibits frequency dependent in-phase and anti-phase oscillations. In: Proceedings of the international joint conference on neural networks 1992, vol IV. IEEE Press, Baltimore, pp 145–151

Currie SN, Stein PSG (1989) Interruptions of fictive scratch motor rhythms by activation of cutaneous flexion reflex afferents in the turtle. J Neurosci 9:488–496

Fowler CA, Saltzman E (1993) Coordination and coarticulation in speech production. Lang Speech 36:171–195

Glass L, Mackey MC (1988) From clocks to chaos. Princeton University Press, Princeton

Gracco VL, Abbs JH (1989) Sensorimotor characteristics of speech motor sequences. Exp Brain Res 75:586–598

Gracco VL, Löfqvist A (1994) Speech motor coordination and control: evidence from lip, jaw, and laryngeal movements. J Neurosci 14:6585–6597

Grossberg S, Pribe C, Cohen MA (1997) Neural control of interlimb oscillations. I. Human bimanual coordination. Biol Cybern 77:131–140

Guevara MR, Glass L, Schrier A (1981) Phase locking, period-doubling bifurcations, and irregular dynamics in periodically stimulated cardiac cells. Science 214:1350–1353

Haken H, Kelso JAS, Bunz H (1985) A theoretical model of phase transitions in human hand movements. Biol Cybern 51:347–356

Jordan MI (1986) Serial order in behavior: a parallel distributed processing approach (Technical Report no. 8604). University of California, Institute for Cognitive Science, San Diego

Jordan MI (1988) Supervised learning and systems with excess degrees of freedom (COINS Technical Report no. 88–27). University of Massachusetts, Amherst

Jordan MI (1990) Motor learning and the degrees of freedom problem. In: Jeannerod M (ed) Attention and performance XIII. Erlbaum, Hillsdale, pp 796–836

Jordan MI (1992) Constrained supervised learning. J Math Psychol 36:396–425

Jordan MI, Rumelhart DE (1992) Forward models: supervised learning with a distal teacher. Cogn Sci 16:307–354

Kawato M (1981) Transient and steady state phase response curves of limit cycle oscillators. J Math Biol 12:13–30

Kay BA, Saltzman EL, Kelso JAS (1991) Steady-state and perturbed rhythmical movements: a dynamical analysis. J Exp Psychol Hum Percept Perform 17:183–197

Kelso JAS, Saltzman EL, Tuller B (1986a) The dynamical perspective on speech production: data and theory. J Phonetics 14:29–59

Kelso JAS, Saltzman EL, Tuller B (1986b) Intentional contents, communicative context, and task dynamics: a reply to the commentators. J Phonetics 14:171–196

Laboissière R, Schwartz J-L, Bailly G (1991) Motor control for speech skills: a connectionist approach. In: Touretzky DS, Elman JL, Sejnowski TJ, Hinton GE (eds) Connectionist models. Proceedings of the 1990 summer school. Morgan Kaufmann, San Mateo, pp 319–327

Lee RG, Stein RB (1981) Resetting of tremor by mechanical perturbations: a comparison of essential tremor and Parkinsonian tremor. Ann Neurol 10:523–531

Lennard P (1985) Afferent perturbations during "monopodal" swimming movements in the turtle: phase-dependent cutaneous modulation and proprioceptive resetting of the locomotor rhythm. J Neurosci 5:1434–1445

Lennard PR, Hermanson JW (1985) Central reflex modulation during locomotion. Trends Neurosci 8:483–486

Löfqvist A (1980) Interarticulator programming in stop production. J Phonetics 8:475–490

Löfqvist A (1991) Proportional timing in speech motor control. J Phonetics 19:343–350

Löfqvist A (1992) Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. Lang Speech 35:15–28

Löfqvist A, Yoshioka H (1980) Laryngeal activity in Swedish obstruent clusters. J Acoust Soc Am 68:792–801

Löfqvist A, Yoshioka H (1981) Interarticulator programming in obstruent production. Phonetica 38:21–34

McGowan RS, Saltzman EL (1995) Incorporating aerodynamic and laryngeal components into task dynamics. J Phonetics 23: 255–269

Munhall K, Löfqvist A, Kelso JAS (1994) Lip-larynx coordination in speech: effects of mechanical perturbations to the lower lip. J Acoust Soc Am 95:3605–3616

Nagashino H, Kelso JAS (1991) Bifurcation of oscillatory solutions in a neural oscillator network model for phase transition. In: Proceedings of the second symposium on nonlinear theory and Its applications (NOLTA '91). Shikanoshima, pp 119–122

Nagashino H, Kelso JAS (1992) Phase transitions in oscillatory neural networks. In: Ruck D (ed) Science of artificial neural networks. SPIE Proc 1710:279–287

Nittrouer S, Munhall K, Kelso JAS, Tuller B, Harris KS (1988) Patterns of interarticulator phasing and their relation to linguistic structure. J Acoust Soc Am 84:1653–1661

Pavlidis T (1973) Biological oscillators: their mathematical analysis. Academic Press, London

Pribe C, Grossberg S, Cohen MA (1997) Neural control of interlimb oscillations. II. Biped and quadruped gaits and bifurcations. Biol Cybern 77:131–140

Saltzman E (1991) The task dynamic model in speech production. In: Peters H, Hulstijn W, Starkweather CW (eds) Speech motor dynamics in stuttering. Elsevier, Amsterdam, pp 37–52

Saltzman E (1992) Biomechanical and haptic factors in the temporal patterning of limb and speech activity. Hum Movement Sci 11:239–251

Saltzman EL (1995) Intergestural timing in speech production: Data and modeling. In: Elenius K, Branderud P (eds) Proceedings of the XIIIth international conference of phonetic sciences (ICPhS 95), vol 2. Arne Strombergs Grafiska, Stockholm, pp 84–91

Saltzman EL, Munhall KG (1989) A dynamical approach to gestural patterning in speech production. Ecol Psychol 1:333–382

Saltzman E, Kay B, Rubin P, Kinsella-Shaw J (1991) Dynamics of intergestural timing. Perilus 14:47–56 (Institute of Linguistics, University of Stockholm, Stockholm)

Saltzman E, Löfqvist A, Kinsella-Shaw J, Rubin P, Kay B (1992) A perturbation study of lip-larynx coordination. In: Ohala JJ, Nearey TM, Derwing BL, Hodge MM, Wiebe GE (eds) Proceedings of the international conference on spoken language processing (ICSLP '92), addendum. Priority Printing, Edmonton, pp 19–22

Saltzman E, Löfqvist A, Kinsella-Shaw J, Kay B, Rubin P (1995) On the dynamics of temporal patterning in speech. In: Bell-Berti F, Raphael L (eds) Studies in speech production: a festschrift for Katherine Safford Harris. American Institute of Physics, Woodbury, pp 469–487

Saltzman E, Löfqvist A, Mitra S (1998) "Clocks" and "glue" – global timing and intergestural cohesion. In: Broe M, Pierrehumbert J (eds) Papers in laboratory phonology V (in press)

Tuller B, Kelso JAS (1991) The production and perception of syllable structure. J Speech Hear Res 34:501–508

Winfree AT (1980) The geometry of biological time. Springer, New York