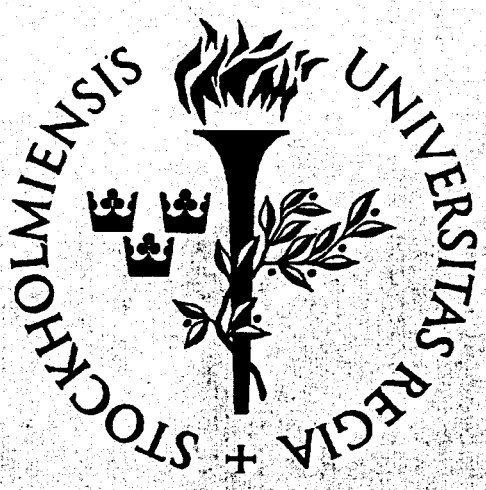
Phonetic
Experimental
Research at the
Institute of
Linguistics
University of
Stockholm



PERILUS XIV
December 1991

Papers from the symposium

Current Phonetic Research Paradigms: Implications for Speech Motor Control

held in Stockholm, August 13 – 16, 1991

Edited by Olle Engstrand and Catharina Kylander

DYNAMICS OF INTERGESTURAL TIMING

Saltzman, E. (1,2), Kay, B. (3), Rubin, P. (1), & Kinsella-Shaw, J. (1,2).

1) Haskins Laboratories, New Haven CT; 2) Center for the Ecological Study of Perception and Action, University of Connecticut, Storrs CT; 3) Department of Psychology, Brown University, Providence RI, U.S.A.

Introduction

This preliminary report addresses the general issue of how to characterize the dynamics that underlie the temporal patterning of speech gestures. The experimental approach adopted was to apply unpredictable mechanical perturbations to the articulatory periphery during speech sequences, and to examine the resultant changes in the temporal or phasing structure of the sequences. Two types of sequences were used: repetitive (/pæpæpæ.../) and discrete (/pəsæpæpl/). The former sequence types were analyzed using both *phase-resetting* (Saltzman, in press) and *transient* techniques, while the latter were analyzed using transient techniques only. The focus of the study was on discerning the relationships between the patterns observed in the repetitive and discrete data, and on identifying common dynamical principles (Saltzman, Kay, Kinsella-Shaw, and Rubin, in preparation).

Repetitive Data: Phase-resetting analyses

Recent data on unimanual oscillatory movements (Kay, 1986; Kay, Saltzman, & Kelso, 1991) have demonstrated that transient mechanical perturbations delivered to the motor periphery induce permanent and systematic shifts in the phasing of such rhythmic movements. Gracco & Abbs (1989) have shown that similar perturbations during single, noncyclic productions of /sæpæpl/ have the effect of inducing systematic shifts in the timing or phasing of subsequent movement elements. However, since these speech data are from short sequences, it is not possible to determine if the perturbation actually resets an underlying sequence "clock", or if the temporal shifts were due to systematic effects in the articulators' transient, post-perturbation behavior. In order to distinguish between these two possibilities, we conducted a *phase-resetting* experiment on the production of extended, repetitive speech sequences .

The goal of phase resetting analyses (e.g., Glass & Mackey, 1988; Kawato, 1981; 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. 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 place a major constraint on theories of speech production that posit a central timing network or "clock" underlying the production of such sequences (e.g., Saltzman & Munhall, 1989). Such a result would imply that the hypothesized central timekeeper could not simply drive the articulatory periphery in a strictly feedforward, unidirectionally coupled manner. Rather, the hypothesized central timer and the peripheral musculoskeletal apparatus must be coupled bidirectionally, so that feedback information concerning the biomechanical state of the periphery can influence the functioning of the timer.

Methods

Subjects and Equipment. In this experiment (also reported in Saltzman [in press]), a single speaker of American English was seated in an adjustable dental chair, with his head restrained in an external frame. A small paddle connected to a torque motor was placed on the lower lip with a tracking force of 3 gm., in order to deliver step pulses of downward force (50

gm.) at random times during the experimental trials. Timing of perturbation onset was controlled by a VAXstation II/GPX. Infrared light-emitting diodes were mounted on the upper lip, lower lip, lip paddle, jaw, nose (the nose LED acted as a spatial reference), and movements were measured optoelectronically. Additionally, the acoustic speech signal and control voltage applied to the torque motor were recorded. All data was fed into a 16 track FM tape recorder for later digitization.

Protocol. Two experimental sessions were conducted, each lasting approximately 3 hours, and 12 blocks of 25 trials were performed per session. Blocks alternated between *repetitive* and *discrete* experimental conditions. In the discrete condition, each trial consisted of the single "word" /pəsæpæpl/; in the repetitive condition, each trial consisted of a sequence of approximately 15-20 repetitions of the syllable /pæ/, spoken at a syllable rate comparable to that used in the discrete trials. Details of the experimental protocol for the repetitive blocks will be described further in the present section; the protocol for the discrete blocks will be described below in the section *Discrete Data: Transient Analyses.*

For the repetitive blocks, perturbations were delivered during a random sampling of 80% of the trials; perturbation duration was preset in an external timing circuit to equal the subject's average syllable duration measured during pretest repetitive trials. On each perturbation trial, the perturbation was delivered during the *nth* syllable (*n* varied randomly from 8-11), and after *m%* of the predetermined syllable duration (*m* varied randomly from 1-100). Task instructions were to not actively resist the perturbation, and to return to a steady rhythm similar to that produced before the perturbation as quickly and easily as possible.

Results

For each perturbation trial, a lip-aperture (LA) trajectory was defined by subtracting the upper lip signal (UL) from the lower lip signal (LL), i.e., LA = LL – UL (Figure 1 illustrates LA trajectories from two perturbation trials). Individual cycles were then defined between successive peak bilabial openings, and four cycle types were identified: a) *pre-perturbation* cycles included the trial's first cycle through the last cycle prior to the one containing the perturbation onset; b) *perturbation* cycles included all cycles that overlapped the perturbation interval; c) *transient* cycles were defined as those cycles following the perturbation during which cycle periods deviated from the average pre-perturbation cycle period by more than an absolute-valued percentage criterion (this criterion was set on a trial-by-trial basis to equal the absolute value of the largest percentage deviation of the pre-perturbation cycles from their own average period.); and d) *post-return* cycles were defined from the last transient cycle to the end of the trial.

Cycle phase, ø, was defined to be zero at all peak bilabial openings. For all other points between peak openings, phase was defined as (t / T_i), where t is the time (in secs) from the most recent peak preceeding a given event of interest, and T_i is the period (in secs) of the cycle containing the event. The phase of perturbation delivery was defined with respect to the time of perturbation offset. This offset time served as a temporal anchoring point for "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 an *average old phase*, $\overline{ø_{old}}$, and *average new phase*, $\overline{ø_{new}}$, respectively. *Phase shift*, Δø, was then defined as $(\overline{ø_{new}} - \overline{ø_{old}})$ (modulo 1). Thus, Δø is the amount that a given trial's post-return rhythm has been shifted relative to its pre-perturbation rhythm (Δø > 0 denotes phase advance; Δø< 0 denotes phase delay). The same measures were obtained for the control (no perturbation) trials, where calculations were anchored to the end of a randomly timed, but not delivered, "perturbation".

Figure 2 illustrates the results of our analyses (one subject, two sessions), using data binned according to intervals of old phase. As can be seen in the figure, the rhythm showed a phase advance in the .2-.4 interval that was significantly different from the no-perturbation

control trials (Dunnett's test, $p<.05$). Thus, these data support the hypothesis that central timing processes for speech are indeed sensitive to appropriately timed mechanical perturbations of the biomechanical periphery, and that such events can permanently reset the rhythms of such central "clocks."

Discrete Data: Transient analyses

Since speech phase-resetting data have not been reported previously in the literature, it is reasonable to ask whether these data are relevant to understanding the production of non-repetitive, discrete speech sequences, e.g., single words. Additionally, since the reported phase-resetting data was collected from a single subject, how generalizable are these data to other speakers, i.e., is the subject a "representative" speaker? The discrete sequence blocks of the present experiment were included to address these questions.

Protocol. In the discrete blocks, each trial consisted of the single "word" /pəsæpæpl/, and perturbations were delivered during a random sampling of 80% of the discrete trials. Perturbation duration was preset to equal the subject's average acoustic interval between the onsets of the first and second /æ/s, measured during a set of pretest discrete trials. Pretest measures were also used to parameterize a random timing circuit for controlling perturbation onset. This circuit was triggered by the acoustic onset of the initial /ə/, and allowed perturbation offsets to occur at m% of the pretest interval between the onsets of the first and second /æ/s (m varied randomly from 1-100). Task instructions were to not actively resist the perturbation, and to continue speaking as naturally as possible.

Results

To date, only data from the first session have been analyzed. As in the phase resetting analyses, lip aperture trajectories were computed for each trial. For all trials, the time intervals from peak aperture during the first /æ/ to peak aperture during the second /æ/ were measured. For the perturbed trials, these intervals were binned and averaged according to the time of perturbation offset minus the time of interval offset, and compared to the mean value of the control (nonperturbation) trials' durations (Figure 3). T-tests (for unequal variances) were made, comparing each of the four perturbation bins with the nonperturbation control bin. In order to adjust for an elevated Type I error rate due to multiple comparisons, we selected criterion α-levels accordingly by dividing .01 and .05 by the number of comparisons made. Results indicated that perturbations induced significant duration shortening when perturbation offsets were either synchronous with (bin 0; $t(25) = 8.937$, $p < .002$) or slightly followed (bin 100; $t(19) = 3.256$, $p < .01$) the duration offset. These findings agree with the data of Gracco & Abbs (1989). However, these authors also showed significant duration lengthening when perturbation offsets were roughly synchronous with the duration onset; our data only showed a nonsignificant tendency to lengthen under these conditions. Nevertheless, we considered our data patterns to be close enough to previously reported patterns to consider our subject a "representative" speaker, and proceeded to consider the relationship between the perturbation-induced temporal changes observed in the repetitive and discrete data.

Repetitive Data: Transient analyses

Transient analyses were conducted for the repetitive data (/pæpæ... /; see the above section *Repetitive Data: Phase-resetting analyses* for data collection protocol), focusing on the durations of the first and second *perturbation cycles*, i.e., the first two cycles that overlapped the perturbation interval. As with the phase-resetting analyses described earlier, control trial values were calculated for first and second "perturbation cycles" that were defined by randomly timed, but not delivered, "perturbations". To date, only data from the first session have been analyzed.

Figure 4 defines in schematic form the temporal landmarks and intervals used in the present analyses. Data for the durations of the first perturbation cycles (dur_1) were normalized according to the average duration of their corresponding preperturbation cycles ($\overline{prepert}$),

using the formula (dur_1 – $\overline{prepert}$) / $\overline{prepert}$. Figure 5 displays these data after binning and averaging according to the normalized time of perturbation onset minus the onset time of the first perturbation cycle, using the formula (perton – t_{on1}) / $\overline{prepert}$. Results indicated that cycle durations were significantly shortened relative to controls for the first three bins, and significantly lengthened for the final bin (Dunnett's test, $p<.01$). However, the lengthening in the final bin appears to have been due largely to two relatively uninteresting factors. The first factor was a passive mechanical consequence of the lip-opening perturbation onset causing a passive mechanical delay of the immediately following onset of lip closing. The second factor was a statistical artifact of the definition of the final bin itself, since it contained "long" cycles in which the perturbation onsets occurred at a time greater than 100% of the corresponding average preperturbation cycle duration. Thus, as the perturbation "slides" back into the first perturbed cycle, there is no effect (ignoring the lengthening in the first bin for the previously stated reasons) until approximately half of the cycle is covered, after which shortening begins.

Data for the second perturbed cycles were similarly normalized, using the formula (dur_2 – $\overline{prepert}$) / $\overline{prepert}$, and displayed in binned and averaged form according to the formula (pertoff – t_{off2}) / $\overline{prepert}$ (Figure 6). Results indicated that cycle durations were significantly shortened when perturbation offsets were either synchronous with (bin 8) or slightly followed (bin 43) the cycle offset (Dunnett's test, $p<.01$). These results are comparable to the results obtained for the single word condition (see the *Discrete Data: Transient analyses* section), i.e., as the perturbation "slides" forward into the cycle, there is no effect until approximately 90% of the cycle is covered and shortening begins.

Relationship between transient and steady-state data

How are the duration changes observed in the perturbation cycles of the repetitive data related to the steady-state, phase-resetting results obtained in these same data? We hypothesized that most of the phase shifts observed in the steady-state were due to duration changes induced during the first two perturbation cycles. In order to test this hypothesis, the summed durations for each trial's first and second perturbed cycles were normalized using the formula ([dur_2 + dur_2] – 2 * $\overline{prepert}$) / $\overline{prepert}$, and displayed in binned and averaged form using the formula (perton – t_{on1}) / $\overline{prepert}$ (Figure 7). Results indicated that the summed perturbation cycle durations were significantly different (shorter) than the corresponding nonperturbation controls only in the first two bins (Dunnett's test, $p<.01$). These trials' data were then replotted using the same abscissa, but with the ordinate representing the steady-state phase shift (Figure 8; see the earlier section *Repetitive Data: Phase-resetting analyses* for the method used to compute phase-shift). Results were comparable to the original phase-shift analyses (Figure 2), showing a significant phase-shift (advance) only at a single data bin. That this effect occurred in different bins in Figure 2 (bin 2) and Figure 8 (bin 1) results from the different conventions used to define the corresponding abscissa values, and from the fact that the perturbation duration was slightly longer than the average preperturbation cycle durations.

Discussion

All the data (repetitive and discrete) appears consistent with the hypothesis that speech production displays both *phase sensitivity* and *temporal saturation* to experimentally induced lip-opening perturbations in the utterance types examined. Phase sensitivity is indicated by the fact that significant effects (phase advances or duration shortenings) were only shown when the perturbations occurred during lip *closing* gestures; temporal saturation was indicated by the fact that phase advances or syllable shortenings were limited to approximately 20-30% of the duration of an unperturbed control syllable. Phase sensitivity is most easily seen in the first perturbation cycle of the repetitive data (Figure 5), for which no shortening occurs until the perturbation encroaches on the cycle's closing phase. Temporal saturation is most

until the perturbation encroaches on the cycle's closing phase. Temporal saturation is most easily seen in the second perturbation cycle of the repetitive data (Figure 6; recall that this condition is comparable to the discrete word data of Figure 3). When the perturbation offset occurs early in the second cycle, there is no effect, since the previous cycle has already shortened up to the systems saturation limit; when the perturbation offset occurs late in the second cycle, the cycle shortens since the previous cycle has been perturbed only during its nonsensitive *opening* phase, thereby allowing the sensitive closing phase of the second cycle to induce shortening.

Summary and Conclusions

Phase-resetting analyses of a single speaker's repetitive speech sequences have shown that mechanical perturbations can induce long-lasting phase shifts, thus supporting the hypothesis of an underlying timing network for speech that is bidirectionally coupled with, yet functionally distinct from, the dynamics of the peripheral articulatory apparatus. Further, analyses of perturbed productions of single words by this same speaker proved comparable to data reported previously in the literature, thereby allowing us to conclude that the repetitive data is not likely to be temporally anomalous or nonrepresentative. Finally, the transient analyses of the repetitive data show that the patterns of temporal change are consistent across both repetitive and discrete sequences, thereby allowing us to conclude that the phase-resetting results are not mere artifacts of the task's cyclicity, but that they reflect the dynamics underlying the production of discrete words as well. More data is needed, however, not only to corroborate these findings but to generalize them beyond a single speaker's production of the relatively simple bilabial sequences that were described in this report.

ACKNOWLEDGMENTS

The work reported in this paper was supported by grants to Haskins Laboratories from the following sources: NIH Grant NS-13617 (Dynamics of Speech Articulation) and NSF Grant BNS-8520709 (Phonetic Structure Using Articulatory Dynamics). The contributions of Patrick Haggard, Scott Kelso, and Kevin Munhall to the collection and analysis of data from several pilot studies are gratefully acknowledged

REFERENCES

Glass, L. G. & Mackey, M. C. (1988). **From clocks to chaos: The rhythms of life.** Princeton, NJ: Princeton University Press.

Gracco, V. L., & Abbs, J. H. (1989). Sensorimotor characteristics of speech motor sequences. **Experimental Brain Research, 75,** 586-598.

Kawato, M. (1981). Transient and steady state phase response curves of limit cycle oscillators. **Journal of Mathematical Biology, 12,** 13-30.

Kay, B. A. (1986). **Dynamic modeling of rhythmic limb movements: Converging on a description of the component oscillators.** Unpublished doctoral dissertation, Department of Psychology, University of Connecticut, Storrs, CT.

Kay, B. A., Saltzman, E. L., & Kelso, J. A. S. (1991). Steady-state and perturbed rhythmical movements: A dynamical analysis. **Journal of Experimental Psychology: Human Perception and Performance, 17,** 183-197.

Saltzman, E. L. (in press). Biomechanical and haptic factors in the temporal patterning of limb and speech activity. **Human Movement Science.**

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

Saltzman, E. L., Kay, B. A., Kinsella-Shaw, J., & Rubin, P. E. (in preparation). Temporal structure in short and extended speech sequences: Perturbation analyses.

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

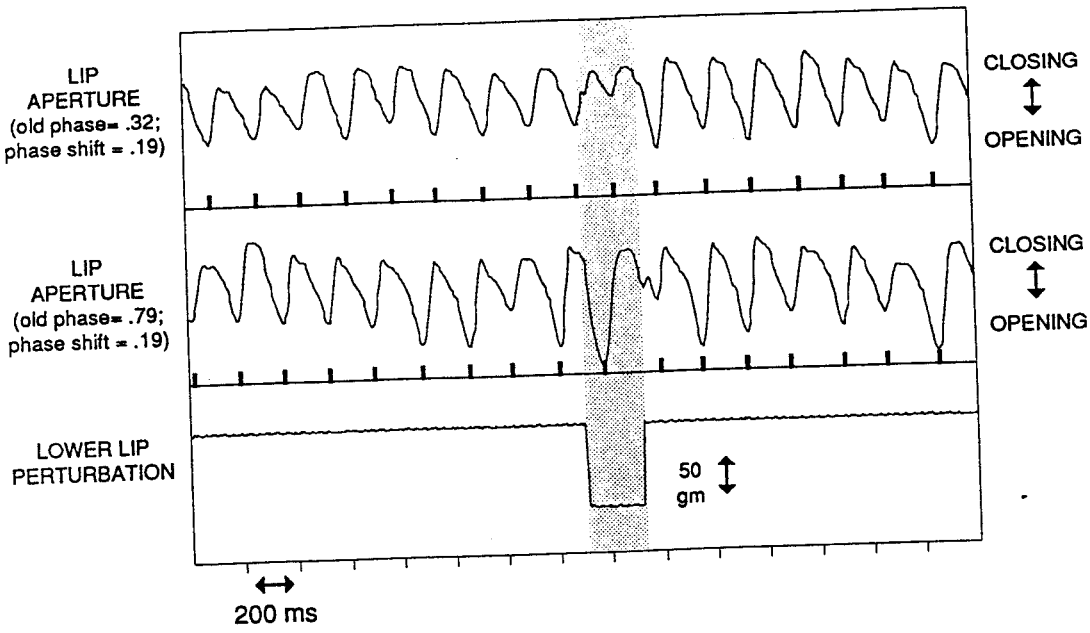


Figure 1. Lip aperture trajectories for two perturbation trials are shown in top two rows. Bottom row shows time-aligned trajectory for perturbation force.

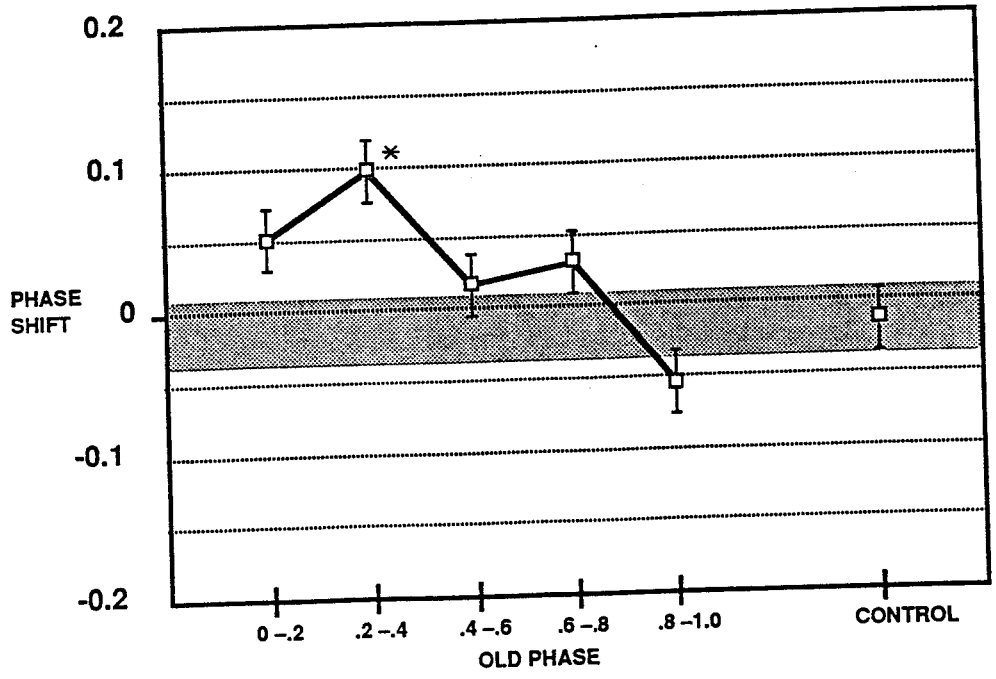


Figure 2. Summary of phase shift data. Number of trials included in bins, from left to right, are 42, 34, 34, 33, 32, 52. (Open squares = means. Vertical bars = standard errors.)

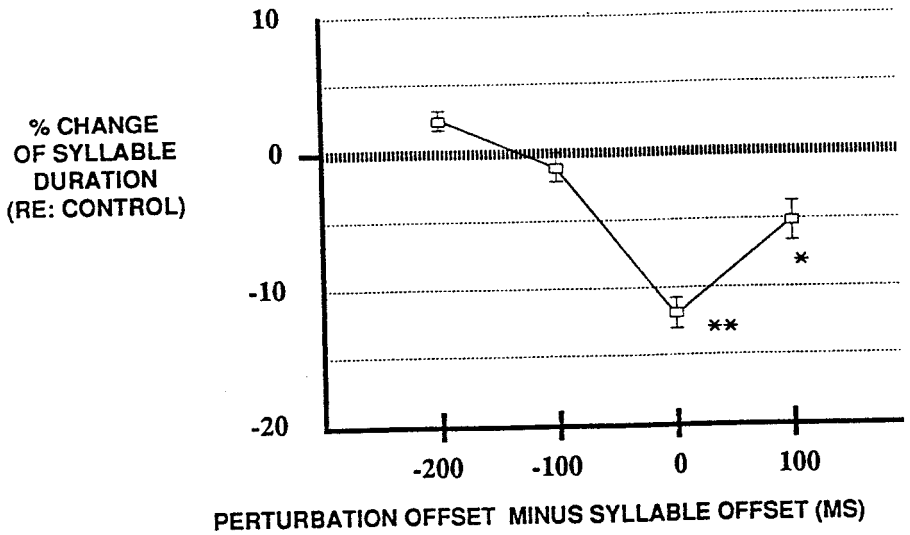


Figure 3. Summary of discrete data. Number of trials included in perturbation bins, from left to right, are 18, 18, 15, 14. Number of control trials =26.

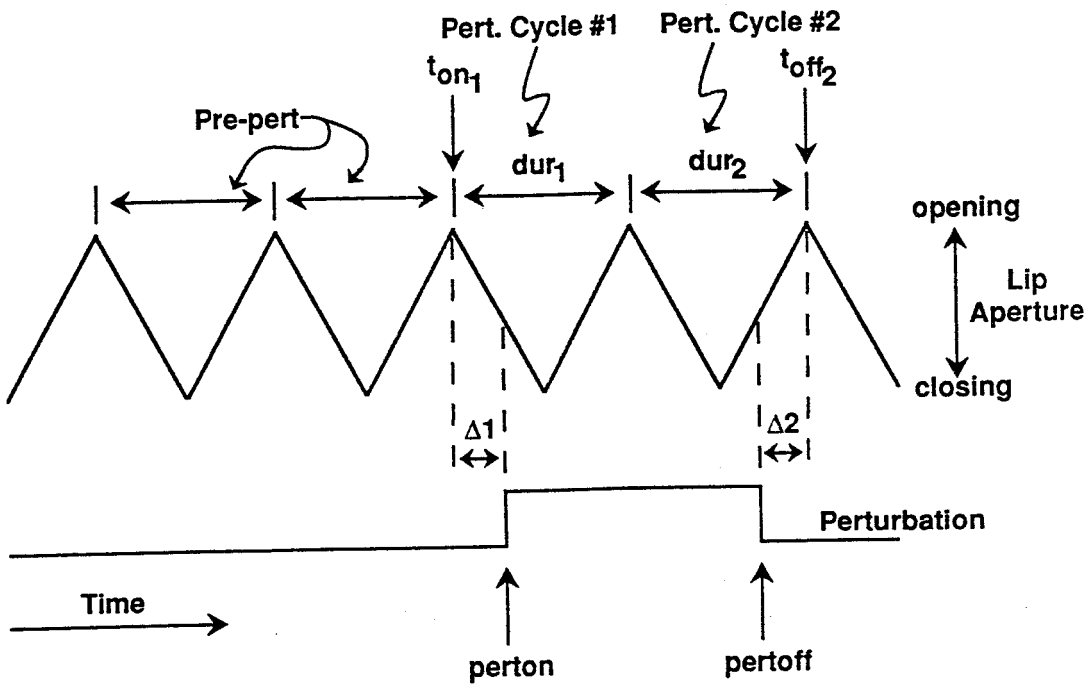


Figure 4. Schematic display of lip aperture and perturbation trajectories for perturbation cycles #1 and #2.

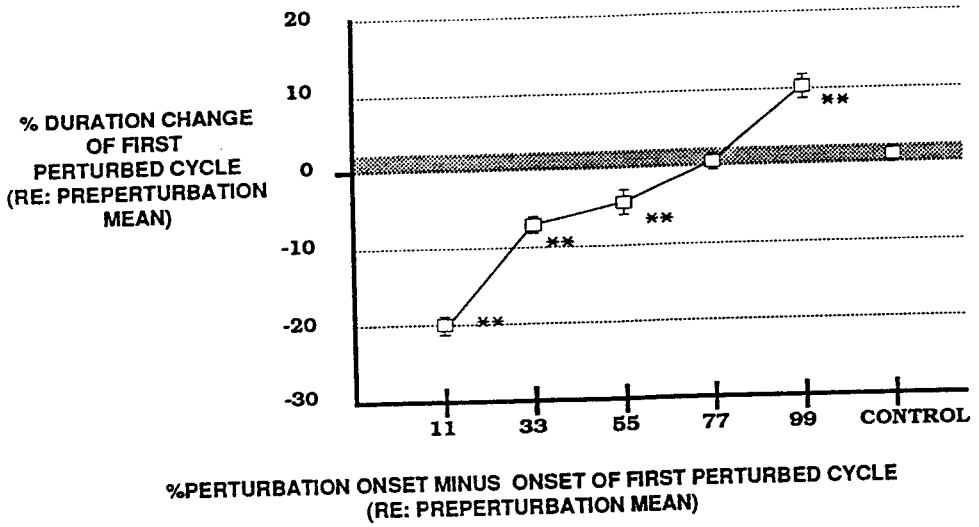


Figure 5. Summary of transient data for first perturbation cycles. Number of trials included in bins, from left to right, are 23, 18, 21, 26, 20, 29.

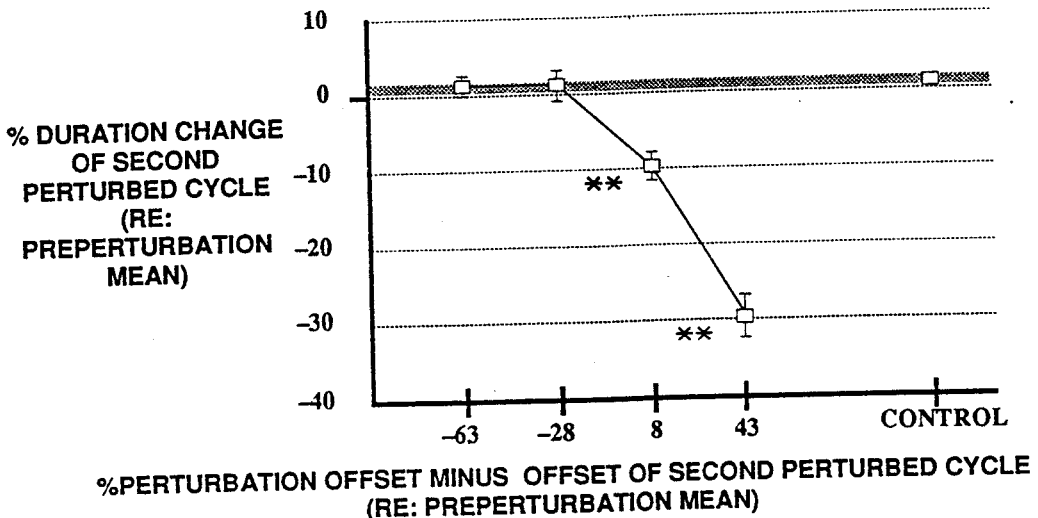


Figure 6. Summary of transient data for second perturbation cycles. Number of trials included in bins, from left to right, are 12, 52, 30, 14, 29.

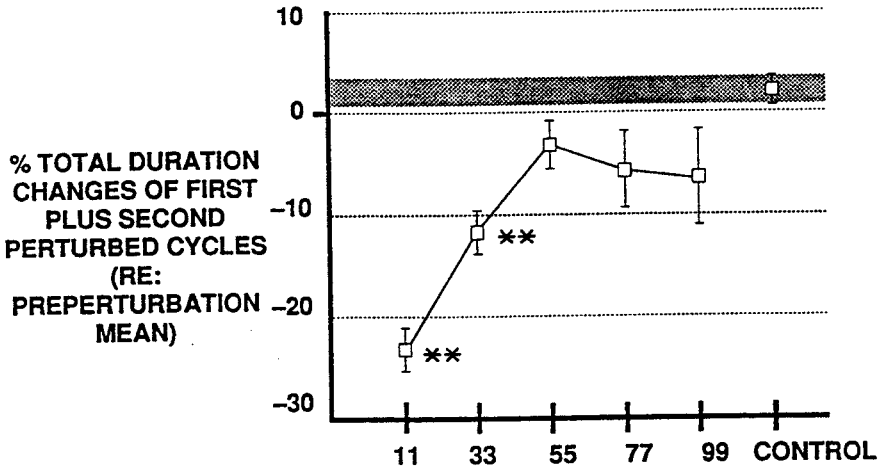


Figure 7. Summary of transient data for first plus second perturbation cycles. Number of trials included in bins, from left to right, are 23, 18, 21, 26, 20, 29.

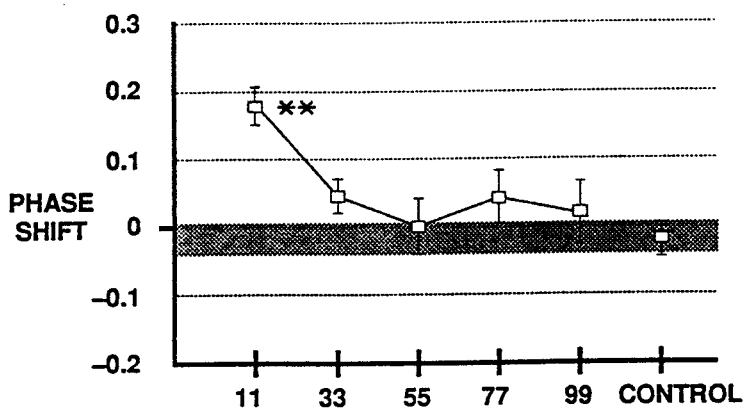


Figure 8. Summary of phase shift data for same trials as shown in Figure 6. Number of trials included in bins, from left to right, are 23, 18, 21, 26, 20, 29.