LONG-TERM SPATIOTEMPORAL STABILITY OF LIP-JAW SYNERGIES FOR BILABIAL CLOSURE

Peter J. Alfonso

The Department of Speech and Hearing Science
The University of Illinois at Urbana-Champaign, U.S.A.

ABSTRACT - Assumptions made about invariant control schemes are gleaned from data collected from a single point in time, that is, data collected from a single session. Thus, we know relatively little about the magnitude of the day-to-day articulatory variability that underlies an invariant percept. The results of the experiment reported here support the notion that the predictable long-term variability in speech motor output that reflects the inherent complexity of the motor system in meeting the time varying demands of rapid conversational speech.

INTRODUCTION

An implicit notion in many dynamical models of movement is that there is an underlying invariance in the motor control scheme despite the observed surface variations in performance (e.g., Saltzman, 1991; Saltzman & Munhall, 1989). However, the assumptions that we make about invariant control schemes are gleaned from data collected from a single point in time, that is, data collected from a single session and usually within a single speech rate (e.g., Gracco & Abbs, 1986; Gracco, 1994). We know far less about the magnitude of the day-to-day articulatory variability that underlies an invariant percept. The overall purpose of the experiment described here is to explore further the saliency of certain presumed invariant motor characteristics of speech, and by extension, the notion of coordinative structures, by comparing certain spatiotemporal characteristics of tongue-jaw and lip-jaw movements for stops and fricatives across multiple sessions and across speech rate.

METHODS

The movements of the tongue blade, lips, and jaw were transduced by a magnetometer system. (electromagnetic midsaggital articulography (EMMA)). EMMA signals were digitized at 400 Hz while the corresponding speech acoustic signals were digitized at 10 kHz. A single session included twenty repetitions of /pap/, /tat/, and /sas/ imbedded in a carrier phrase at normal, slow, and fast speech rates. The 60 phrases were blocked by rate and produced first at a normal speech rate, than again at a fast rate, and finally at a slow rate. Sessions were repeated three times, and the interval between sessions was about two weeks. Displacement onset and offset for syllable initial stop and fricative closure were marked by standard velocity criteria. Considerable software development, calibration procedures, and hardware modifications were made to the Carstens EMMA system used here and in collaboration with Phil Hoole at the University of Munich and Pascal van Lieshout at the University of Nijmegen. A summary of some of these modifications can be found in Alfonso, et. al., (1993) and Van Lieshout, et. al., (1994).

RESULTS

General trends in the organizational stability for /p/ closure are demonstrated in Figure 1. Note that organizational patterns differ across subjects. Subject 6 achieves closure primarily by jaw and upper lip displacement, the dark and hashed sections of the bars, the lower lip contributing the least. On the other hand, Subject 5 achieves closure predominately by jaw displacement with relatively less contribution of the lips, especially the upper lip. Next, note that the relative displacement patterns are generally stable across sessions for 4 of the 7 subjects. As a first approximation of across session stability, inconsistencies across sessions in the primary articulator, that is, the articulator that contributed the most toward closure, were noted. These inconsistencies are noted by the arrows in Figure 1, which shows a change in control strategies

for subjects 1, 3, and 7. For example, in the case of Subject 3, /p/ closure is achieved primarily by the jaw in session 1 but by the lower lip in sessions 2 and 3. Subjects 1 and 3 show a shift in the predominate articulator from the jaw in session 1 to equal contribution of the jaw and upper lip in session 2.

While complete pattern analyses might yield different interpretations, for example, the patterns appear quite different for Subject 2 although the jaw is the consistent primary articulator, Figure 1 shows that about 3 of the 7 subjects appear to enlist different control strategies across sessions. In absolute terms, these differences can be large. For Subject 2, for example, average vertical displacements of the non-dominate upper and lower lips vary 200 percent across sessions, from about 2 to 4 mm, while total displacement of the jaw and lips synergy vary from 10 to 14 mm across sessions.

However, Figure 2 shows that while the control strategies can and do vary across sessions for some subjects, the movements of the lips and jaw are well coordinated in each of the sessions and for all subjects. Shown in Figure 2 is the coefficient of variation for the upper lip in open bars, the lower lip in gray bars, the jaw in dark bars, and the total displacement in hashed bars across sessions and subjects. The figure shows that motor equivalence covariability is stable even in the case of subjects 1, 3, and 7 who demonstrated instability across sessions, and even in the case of subject 2, the only subject who demonstrated unstable spatial and temporal organizational patterns for both labial and alveolar gestures. Only in the case of the third session for Subject 6 is motor equivalence covariability not observed.

Turning next to temporal stability, Figure 3 shows the relative distribution of the upper lip, lower lip, and jaw sequence patterns for /p/ closure. The dark portions of the bars represent the upper lip, lower lip, and jaw sequence, the light portions represent the lower lip, upper lip, and jaw sequence, and the hashed portions represent all others. By far, the predominant sequences are those in which lip movement occurs first and jaw movement occurs last. First, the figure shows that either lip lead sequence is equally likely to occur. For example, Subject 1 prefers the lower lip lead sequence while Subject 3 prefers the upper lip lead sequence. Second, the figure shows that some subjects, for example Subjects 6 and 7, show no clear preference for either lip lead sequence. Third, the figure shows that two of the subjects, 2 and 4, show a clear reversal in the lip lead sequence across sessions.

One of the reasons that the sequence pattern is not stable across time is that it does not reflect interarticulator relative time. Subjects who demonstrate tight coupling of lip movements, for example, would have a higher probability of producing both lip-lead sequences compared to subjects who demonstrate longer relative timing of the lip movements. This is demonstrated in Figure 4, which shows the lip relative time and the lagging lip to jaw relative time for the session average sequences. A session average sequence represents the ensemble averaged trajectories associated with the movements of the lips and jaw. Note that the relative time for lip movements in the case of Subjects 6 and 7 is less than five ms. Recall that Figure 3 showed that these subjects demonstrate nearly equal probability of producing either lip-lead sequences. Thus, a temporal invariancy criterion based on consistent sequence patterns would exclude subjects 6 and 7 whereas a criterion based on tight interarticulator timing would include the same subjects.

Figure 5 graphs both the sequence pattern and relative time and shows that three of the subjects are inconsistent in at least one of these two temporal parameters across sessions. The hashed bars represent the lower lip, upper lip, and jaw sequence and the light-dark solid bars represent the upper lip, lower lip, and jaw sequence. Subject 2 shows a sequence reversal and Subjects 1 and 4 show relatively large differences in interarticulator relative time across sessions.

DISCUSSION

The results thus far indicate that traditional spatial organizational characteristics, such as displacement and velocity profiles, and that traditional temporal organizational characteristics, such as temporal order and interarticulator relative time, are not stable across sessions for some

subjects. Only motor equivalence covariability is consistently correlated with the invariant percept and even in the instances of unstable spatial and temporal characteristics.

One of the interesting questions to arise from these data is the nature of the precursors to the differing control strategies demonstrated by some of the subjects. Speech articulation is inherently variable due to a large number of time varying demands upon the speech motor control system such as speech rate and articulatory precision. With regard to the relationship between speech rate and spatiotemporal stability; In sessions 2 and 3, subjects were asked to model the rate that they self selected in session 1. In most cases, differences in phrase durations were relatively small at about 15 percent. The relationship between rate variation and spatiotemporal stability is not straightforward and does not appear to account for all of the across session instability. For example, Subject 2, who consistently demonstrates varying control strategies in /t/ and /s/ closure in addition to that shown here, was very successful in maintaining a similar rate across sessions. With regard to articulatory precision, we are assessing other kinematic measures to estimate better the precision in which articulatory gestures are achieved, and how target accuracy bears on the spatiotemporal instability shown here.

In conclusion, the results support the notion that there is an inherent long-term variability in speech motor output that reflects the complexity of the motor system in meeting the time varying demands of rapid conversational speech. Further, the observed stability in the coordination index, namely, motor equivalence covariability, even in the case of instability of the displacement and velocity profiles, temporal order, and interarticulator relative timing are in agreement with the idea that spatial and temporal organization of functionally linked articulators is secondary to gestural specification. This is precisely what the coordinative structure notion within a task-dynamic framework would predict; that the spatial and temporal organization of the articulators that comprise an articulatory complex represent the natural consequence of gestural specification and therefore would not demonstrate stability across sessions.

REFERENCES

Alfonso, P.J., Neely, J.R., van Lieshout, P.H., Hulstijn, W., & Peters, H.F. (1993). *Calibration, Validation, and Hardware-Software Modifications to the Carstens EMMA System.* Proceedings of the ACCOR Workshop on Electromagnetic Articulography in Phonetic Research. (Forschungsberichte 31, pp.105-120). Munich, Germany: Institut fur Phonetik und Sprachliche Kommunikation der Universitat Munchen (FIPKM). Also in Haskins Laboratories Status Report of Speech Research, SR-114, 101-112.

Gracco, V.L. (1994). *Some organizational characteristics of speech movement control.* Journal of Speech and Hearing Research, 37, 4-27.

Gracco, V.L. and Abbs, J.H. (1986). *Variant and invariant characteristics of speech movements.* Experimental Brain Research, 65, 156-166.

Saltzman, E. (1991). *The task dynamic model in speech production.* In H.F.M. Peters, W. Hulstijn, and C.W. starkweather (Eds). Speech motor control and stuttering (pp. 37-52). Amsterdam: Excerpta Medica.

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

Van Lieshout, P.H.H.M., Alfonso, P.J, Hulstijn,W., & Peters, H.F.M. (1994). *Electromagnetic Midsagittal Articulography (EMMA).* In F. J. Maarse, A.E. Akkerman, A.N. Brand, L.J.M. Mulder, and M.J. van der Stelt (Eds.), Computers in Psychology 5: Applications, Methods, and Instrumentation. (pp. 62-76). Lisse, The Netherlands: Swets & Zeitlinger.

ACKNOWLEDGEMENT

Research Supported in part by a Fulbright Research Award and by a University of Illinois Research Board Grant to the first author, and by NIH Grant DC-00121 to Haskins Laboratories.

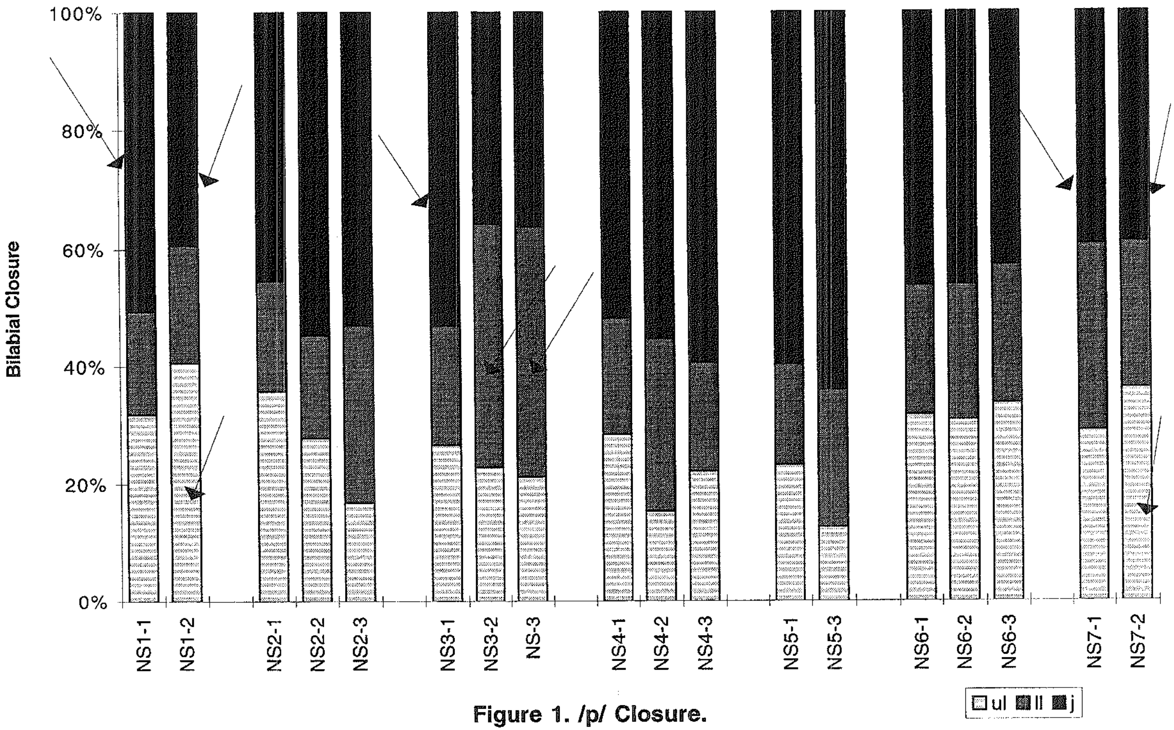

Figure 1. /p/ Closure.
Normalized Vertical Displacement Across Sessions.
Change in Primary Articulator Across Sessions Indicated by Arrow.

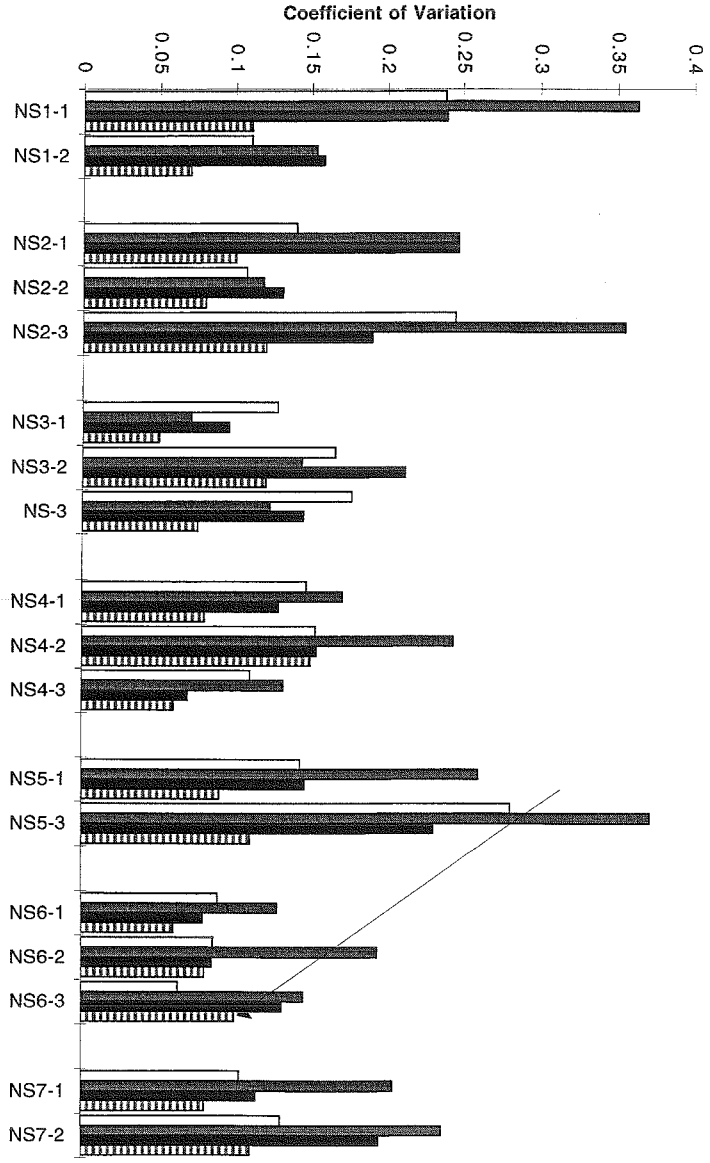

Figure 2. /p/ Closure. Coefficient of Variation.

Synergies that do no meet motor equivalence covariability are indicated by arrows.