2aSC11. Phase transitions in a repetitive speech task as gestural recomposition.

Kenneth de Jong, Byung-jin Lim, and Kyoko Nagao

(kdejong@indiana.edu, bylim@indiana.edu, knagao@indiana.edu)

Department of Linguistics, Indiana University

Main Points

* This paper examines glottal movement data in rate-controlled repetitions of CV (onset consonant + vowel) and VC (vowel + coda consonant) structures.

* It replicates Tuller & Kelso's (1991) observation of rate-induced phase transitions in glottal behavior.

For voiceless consonants:

- CV's are similar at fast and slow rates
- VC's at slow rates do not look like CV's
- BUT: VC's at fast rates change to look like CV's

* **However …** Voiced CV's show the same shifting pattern as voiceless VC's

* **And …** apparent phase transitions may involve both the timing and magnitude of the gestures as well what gestures are employed. VC phase transitions involve a change in gestural composition, rather than just a re-coordination of the same gestures.

Background

OBSERVATION: Stetson (1951 and much earlier):

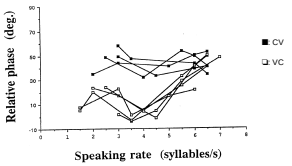
* Early articulatory work
* Repeated VC forms (such as 'eeb') at fast rates
* Perceived as CV (such as 'bee')

MODEL: Modelled by Tuller & Kelso (1991)

* Modelled behavior in terms of a bistable dynamical system
* Stability is indexed by a collective variable: relative phase

The Timing of Peak Glottal Opening expressed in terms of Proportional Time of Successive Oral Closings

* Production study of /p/ shows
  -> CV's have fixed timing at 40 deg. after oral closure regardless of repetition rate
  -> VC's show a transition from earlier phasing in slow rates to the later CV timine in fast rates



ADDITIONS: de Jong (2001a, 2001b):

* Acoustic production studies
* Repeated codas become similar to onsets at fast rates

And

* Fast rate onsets and codas are not neutralized. They retain original differences
* Hence, CV and VC are systematically distinguished by gestural composition in addition to coordination differences

Also

* Rate scaling affects global temporal structure of VC's and CV's in different ways
* Phonemic voicing of the consonant restricts how VC temporal structure is changed
* Hence, while glottal to oral phasing may index CV-VC differences, there is much more to the coordinative differences between the two structures

In the CURRENT STUDY, we …

* Examine glottal actions and timing in rate-varied repetitions
* Actions are are observed via glottal transillumination
* Transillumination traces are parsed with respect to acoustics in order to take a step toward determining why modes in glottal-to-oral phasing are observed

Methods

Stimuli

METRONOME PACER: Repetition rate controlled:
* Two conditions
- start slow (450 ms/syll.), and increase throughout trial (to 200 ms/syll.)
- start fast (200 ms/syll), and decrease throughout trial (to 450 ms/syll.)

* Production rates range across stable production range from 2Hz to 5Hz as found in Nelson et al. (1984)
* Results are qualitatively similar for two conditions, so we will collapse across them for the current paper

TEXT PROMPTS: Simple bisegmental forms:
* Four factors

1) sequence	CV 'bee'	vs.	VC 'eeb'	
2) vowel	/ i / 'eeb'	vs.	/ æ / 'ab'	
3) voicing	'voiced' 'eeb'	vs.	'voiceless' 'eep'	
4) c-type	labial 'pea'	vs.	coronal 'tea'	vs. dorsal 'key'

Total forms = (2*2*2*3) = 24

TOTAL CORPUS per Speaker
24 types * 2 pacers * 3 repetitions = 148 utterances
30 syllables/utterance * 148 utt = 4440 syllables
(analyses still in progress)

Speakers

* 3 Speakers of varied linguistic experience
  - native American English speaking male in 30's
  - non-native Japanese speaking male in 30's
  - non-native Arabic speaking female in 20's
* Speaker patterns are quite different, apparently due to language background
* Current paper focuses on American English speaker

Recordings

* Recorded at Haskins Laboratories
* Include following information:
  - Acoustic traces digitized at 20 kHz
  - Glottal transillumination traces at 635 Hz

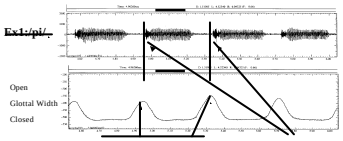
Glottal transillumination traces are the output of a photo-transducer placed externally on the anterior surface of the trachea. This transducer detects a light source placed in the upper pharynx, modulated by the size of the glottis (Baer, et al., 1983). The signals produced by such a system are uncalibrated with respect to the glottal area.

* Also includes the following, not reported on here:
  - EMA trajectories of tongue, jaw, and lips
  - Pharyngeal pressure traces

Results

Patterns of Glottal Movement: Modes

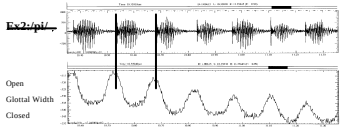
All 'voiceless' CV-forms: SLOW (3 syll/sec)



Glottal abduction movements against a baseline of positioning for voicing.

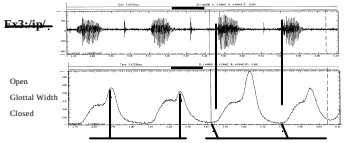
Peak abduction aligned very closely with acoustic release of the consonant

CV-forms: FAST (6 syll/sec)



Roughly the same movement patterns

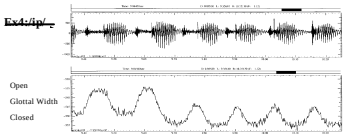
Majority of VC-forms: SLOW (3 syll/sec)



Fast abduction movements (like CV's)

But also slower adductions centered at onset of the vowel.

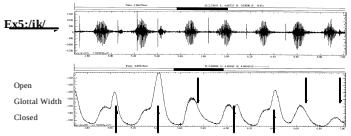
VC-forms: FAST (6 syll/sec)



Very similar to CV forms at fast rate

Irregular abduction in VC forms

Fast abduction movements for consonant devoicing often do not occur, even at slow rates. Hence, some peaks in the signal indicate positioning for voicing in the context of glottal closure. **VC-forms: SLOW (3 syll/sec)**

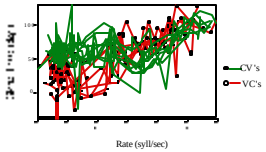


Here we have four with devoicing abductions, and three without.

Phase Histories

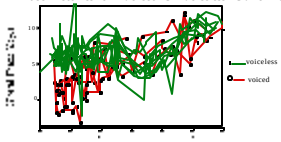
We extracted the timing of the devoicing abduction peaks (not peaks due to voicing in the context of glottal closure). Below, the phasing of peak glottal abduction relative to mid-points of acoustic closure is plotted against syllable-duration

Phase Histories for 'Voiceless' CV-forms & VC-forms



Phase transitions similar to Tuller & Kelso (1991)
* Phasing is roughly fixed throughout rates for CV's
* Phasing is somewhat earlier for slow VC's
* Fast VC's have later phasing, similar to CV's
* Magnitude of the effect is a bit smaller than T+K's

Phase Histories for 'Voiced' & 'Voiceless' CV-forms



'Voiced' consonants in CV's also show phase shifting

Acknowledgements

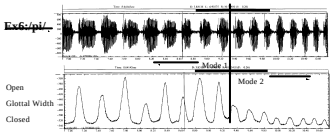
Work supported by the NIDCD (grant #R03 DC04095-01A2) and by the NSF (grant #BCS9910701)

We also express deep appreciation to Anders Lofqvist, without whose seemingly endless patience and remarkable expertise, none of this work cou ld have been done. We also acknowledge Mafuyu Kitahara and Bushra Zawaydeh for their help on this project.

Mode Shifting Patterns

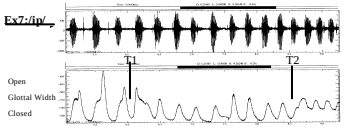
In addition to phase shifts, there are other modal changes induced by rate increases (3.5 - 5.5 syll/sec)

CV's: Non-gradient scaling of magnitude



In half of CV's, there are suggestions of modal behavior, even though the temporal phasing is consistent

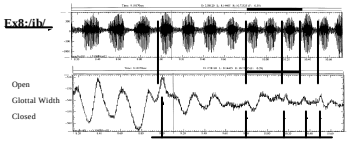
VC's: Loss of glottal adduction



After T1: Intermediate hybrid forms have transposition of adduction movement and abduction peak,

After T2: sudden loss of adduction movement altogether

Voiced VC's: Introduction of glottal abduction



Voiced VC forms show clear addition of abduction movements synchronized with consonant release

Discussion & Summary

* We find modal patterns of glottal behavior which are rate sensitive (as per Stetson, 1951).
* Glottal-to-oral phasing shows transitions corresponding to these modes (as per Tuller & Kelso, 1991)

However:

* Rate sensitive shifts in phasing are not uniquely indicative of syllabic organization: voiceless unaspirated onsets (e.g. /bi/) behave like voiceless codas (e.g. /ip/)
* Gestural magnitude also is involved in modal shifting
  -> Similarity of voiceless aspirated onsets and voiceless codas may be due to scale of glottal gestures
  -> Even with fixed phasing, magnitude exhibits modes
* Coda shifting involves more than rephasing, rather adding and losing gestures
  -> Loss of glottalization gestures
  -> Addition of abduction gestures in 'voiced' cases

References Cited

BAER, T., A. LOFQVIST, and N.S. MCGARR (1983). Laryngeal Vibrations: a Comparison between High-speed Filming and Glottographic Techniques. *Journal of the Acoustical Society of America*, **73**: 1304 - 1308.

DE JONG, K.J. (2001a). Rate-induced Resyllabification Revisited. *Language and Speech*, **44**: 197 - 216.

DE JONG, K.J. (2001b). Effects of Syllable Affiliation and Consonant Voicing on Temporal Adjustment in a Repetitive Speech Production Task. *Journal of Speech, Language, and Hearing Research*, **44**: 826 - 840.

NELSON, W.L., PERKELL, J.S., & WESTBURY, J.R. (1984). Mandible movements during increasingly rapid articulations of single syllables: Preliminary observations. *Journal of the Acoustical Society of America*, **75**: 945 – 951

STETSON, R.H. (1951). *Motor Phonetics*. Amsterdam: North-Holland

TULLER, B., and KELSO, J.A.S. (1991). The Production and Perception of Syllable Structure. *Journal of Speech and Hearing Research*, **34**: 501-508