Beyond Simplex and Complex: Gestural coordination in L2 English

Feng-fan Hsieh and Yueh-chin Chang
National Tsing Hua University

**Introduction:** The aim of this study is to investigate if C-center-like timing patterns can be acquired by Taiwanese Mandarin (TM)-speaking L2 learners of English. It is uncontroversial that there are no consonant clusters in Mandarin (putting aside Cw or Cj for now) and consonant clusters are predominantly repaired by epenthesis in loanword adaptation, e.g., brad -> bu.lai.de 'personal name'. But little is known about the actual gestural timing patterns in L2 speech, although due to the well-established Interference Effect of L1, it is expected that stop-liquid clusters in *play*, for example, would be often, if not always, rendered as a disyllable, e.g., /*pəlay*/, or, a "sesquisyllable," e.g., /*pᵊlay*/ by TM-speaking L2 learners of English.

**Method:** Three native speakers of TM and one native speaker of American English (all male in their 20s) participated in this experiment. Judging from the fact that all TM speakers have never been to an English-speaking country and received regular English education of 7 years in Taiwan, we simply assume that these L2 speakers of English are more or less homogenous, at least articulation-wise, as there is no established norm on speech production measures.

**Data analysis:** Kinematic data were collected using Electromagnetic Articulography (NDI Wave). Following Goldstein et al. (2009), among others, the gesture timing was compared in Stop-V, Liquid-V, and Stop-liquid-V contexts to investigate the phonetic expression of (sub-)syllabic organization. Gesture start is labelled at 20% of peak velocity to target (e.g., Shaw et al. 2009). Each token was embedded in the carrier phrase, "___, I say ___ again," and repeated 10 times in a random order. Mean relative timing of gestures to a heterosyllabic consonantal anchor /g/ was calculated using the lag between a gesture start and /g/. The following triads were recorded and analyzed with the help of Mark Tiede's MView: {pay, lay, play}, {pay, ray, pray}, {bow, low, blow}. {bay, ray, bray}, {kay, lay, clay}, {kye, rye, cry}, {gay, lay, glay}, and { gay, ray, gray}. By comparing a given subject's mean lags for stops/liquids with stop-liquid sequences, we calculated the subject's leftward shifts (C vs. CL: *C-shift*) and rightward shifts (L vs. (C)L: *L-shift*) associated with /bl-/, /br-/, /pl-/, /pr-/, /kl-/, /kr-/, /gl-/, and /gr-/. The analytical heuristics are schematized in (1) (where [ ] = lag between stop/liquid gestures to an anchor point; ">" means significantly longer in duration.

| (1) a. Complex: C-center | b. Simplex: C.CV | c. No shift | d. Rightward-only |
|---|---|---|---|
| [play] > [pay]<br>[lay] > p[lay] | [play] > [pay]<br>[lay] = p[lay] | [play] = [pay]<br>[lay] = p[lay] | [play] = [pay]<br>[lay] > p[lay] |

We used the *lme4* package in R (Bates et al. 2012) to perform a linear mixed effects analysis of the temporal differences within the triads. More detail will be presented in the conference.

**Result:** The C-center timing (1a) is consistently found in the data from the English speaker, as expected. Interestingly enough, the anticipated simplex onset in L2 English is *absent* across the board. More specifically, regarding gestural coordination patterns, *play*, for example, is <u>not</u> realized as *p.lay* in the present study (as in (1b); see, e.g., the case of Moroccan Arabic reported in Shaw et al. 2009). The results are summarized in (2):

| | |
|---|---|
| (2) a. C-center timing (=1a) | pray, blow, bray, cry, glay, gray |
| b. Rightward-only (=1d) | *play, clay* |

Finally, the L2 data were further transcribed by a phonetically trained linguist to see whether there is an epenthetic or transient schwa-like vowel. The results of the tallies are shown in (3). Recall that each stimulus was repeated 10 times in the experiment.

| (2) | *play* | *clay* | pray | blow | bray | cry | glay | gray |
|---|---|---|---|---|---|---|---|---|
| Subj 1 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 |
| Subj 2 | 5 | 3 | 0 | 4 | 2 | 0 | 8 | 0 |
| Subj 3 | 0 | 4 | 0 | 2 | 0 | 0 | 5 | 0 |

**Discussion:** Impressionistically speaking, our participants all speak English with a noticeable accent. Nevertheless, the results of our EMA study show that the C-center timing (1a) is successfully acquired in most cases in (2), suggesting that articulatory timing may not be the most likely source of foreign accent. More importantly, the present results further reveal an unprecedented gestural coordination pattern: (1d), as has been found in {pay, lay, play} and {kay, lay, clay}. It has been repeatedly reported in the literature (see, e.g., Goldstein and Pouplier 2014) that consonant clusters may exhibit either the C-center timing/Complex onset (1a) or the simplex onset (1b). The other two logical possibilities, "No shift" in (1c) and Rightward shift-only pattern in (1d), are often ignored. The reason is simple: given that C-V timing is by and large constant, the patterns in (1c-d) are predicted to be self-contradictory, hence impossible. Upon further scrutiny, however, we found that {pay, lay, play} and {kay, lay, clay} do have something in common. First, there must be an aspirated stop {$p^h$ or $k^h$}. Second, the second member of the clusters is the lateral liquid /l/. Remarkably, these two properties are not shared with the stimuli in the triads in (2a).

We suspect that the eccentric timing as in (1d) is probably due to the aspirated stop onset. More precisely, the reason why [pay] = [play] or [kay] = [clay] is because in [$p^hay$], the lag between *p* and the anchor point is substantially lengthened by aspiration in L2 production (but not in L1 production). Or, the start of *p* or *k* is pushed more leftward if the onset is aspirated in an CV syllable. In other words, it is not impossible to see that aspiration (or, longer VOT) induces eccentric timing in gestural coordination, especially when aspiration is phonemic in Mandarin Chinese. Finally, it remains to be seen why (1d) is not found in {pay, ray, pray} and {kye, rye, cry}. Our data show that the r-sound all involved Tongue Tip (TT)-raising, indicating that no bunched /r/ in L2 production, at least in the present study. In contrast, the lateral liquid involves both TT and Tongue Dorsum-raising (e.g., Goldstein et al. 2009). So it may well be the case that the distinct gesture compositions of the two liquid sounds play a role in this regard.
**Conclusion:** In this work, our principal findings are as follows. (i) the C-center timing may be successfully acquired by Mandarin-speaking L2 learners of English, (ii) Our results further show that aspiration (which is phonemic in Mandarin) may have a bearing on eccentric timing of CV. Precisely, aspiration may push the onset more off the nucleus vowel, (iii) one possible consequence for (ii) is that [pay] = [play], as shown in (1d), exhibiting an uncommon gestural timing pattern. The present study also contributes to a growing body of research on language-specific gestural timing relations.

**References:**


Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. *Frontiers in Phonetics and Speech Science*. Beijing, Commercial Press, pp. 239-250.

Goldstein, L., & Pouplier, M. (2014). The temporal organization of speech. In Victor Ferreira, Matthew Goldrick and Michele Miozzo (eds.), *The Oxford Handbook of Language Production*. Oxford: Oxford University Press, pp. 210-227.

Shaw, J., Gafos, A. I., Hoole, P., & Zeroual, C. (2009). Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in articulation. *Phonology*, 26(1), 187-215.