SPEECH ERRORS IN TAIWANESE: AN EMMA STUDY

Feng-fan Hsieh[a], *Yueh-chin Chang*[a], *Wei-rong Chen*[a], *Man Gao*[b], *Christine Mooshammer*[c], *Hosung Nam*[c], *Mark Tiede*[c,d] *& Louis Goldstein*[c,e]

[a]National Tsing Hua University, Taiwan; [b]Dalarna University, Sweden;
[c]Haskins Laboratories, USA; [d]MIT Research Lab of Electronics, USA;
[e]University of Southern California, USA
ffhsieh@mx.nthu.edu.tw

ABSTRACT

Speech errors in Taiwanese are investigated by means of a speeded repetition task. Our results show that the intrusion bias is also attested for word pairs with mismatched onsets, whereas in the alternating coda condition, reduction errors are the most frequent error type. This cross-linguistic difference can be attributed to language-specific implementation of stop codas.

Keywords: speech errors, Taiwanese, gestures, stop codas, checked syllables

1. INTRODUCTION

The purpose of this study is two-fold. The first goal is to test if results similar to those obtained by [2] could be observed for Taiwanese (Southern Min) by replicating their English speech error experiments. [2] used English disyllabic phrases comprised of CVC syllables with alternating onsets (e.g. *cop top*) to elicit data. Comparable stimuli were also used in this work, but with a crucial distinction: in (South) East Asian languages, checked syllables are *not* equivalent to closed syllables in English, even though both may refer to syllables ending with a (voiceless) stop coda (CVS(top)). Regarding checked syllables, stop codas are *never* released, and are consistently reinforced by various degrees of "glottalization" (see [1] for results of a recent laryngoscopic study and references cited therein). By contrast, English CVS syllables do not have such properties (at least in careful/slow speech). Therefore, our first research question is whether or not language-specific implementation of phonetic details plays a role in this regard. We conducted an instrumental investigation of speech errors for checked syllables to see if the gestural intrusion bias and the attested error patterns could also be found in Taiwanese CVS CVS sequences.

The second goal is to investigate an understudied topic, namely, errors in coda positions, by using CVS syllables with alternating codas (e.g. *pop pot*) [3]. Our research questions are: i) whether the intrusion bias could also be found in coda positions [3], and ii) whether language-particular factors (e.g. the fact that stop codas are never released in Taiwanese) affect error patterns.

2. METHODS

2.1. Data collection

The data of this study were collected at Haskins Laboratories, with an electro-magnetic midsagittal articulometer EMMA, [4], which allows the tracking of the kinematic movements of sensors placed on the articulators (e.g. tongue and lips). The kinematic data were collected with a sampling rate of 500 Hz and acoustic data were recorded at 41 kHz.

Five (5) native Taiwanese speakers, aged from mid-20s to mid-30s, were recruited/paid for their participation in this study. No hearing or speech impairments were reported. The data from the only female speaker were excluded from analysis due to some technical problems.

A speeded repetition task was employed to elicit speech errors while collecting the articulatory movement data with an EMMA [2, 3, 5]. The participants were asked to read randomized stimuli (written in Chinese characters) from a computer screen at an increasing rate controlled by an accelerating metronome. Each trial lasted 15 seconds and was repeated four times.

2.2. Stimuli

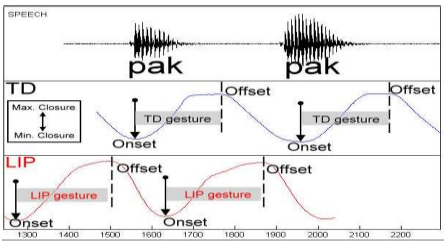
The stimuli are disyllabic phrases containing SVS syllables (S = unaspirated voiceless stop) and the three groups of stimuli are: onset-alternating, coda-alternating, and control/non-alternating phrases. For the onset-alternating ones, we have *kap tap*, *pak tak*, *tak pak*, and *tap kap*, and *pak pat*, *pat pak*, *tak tap*, and *tap tak* were used for the coda-alternating ones, while non-alternating phrases include *kap kap*, *pat pat*, *tak tak*, *tap tap*, and *pak pak*. There are two possible tone patterns: H-L and L-H. Each stimulus was produced with each tonal pattern (cf. different stress patterns in [2, 5]).

2.3. Data analysis

The data were subsequently transformed to MATLAB data files and displayed for checking the data in MView, a MATLAB script developed at Haskins Laboratories. Automatic gesture boundary marking and error counting were conducted with the help of the Visual Basic programs developed by the NTHU Phonetics Lab.

We follow [2] in this work, assuming that a gesture reaches its target when the articulator forms the maximal constriction (or, gesture offsets) and is released afterwards. Gesture onsets, on the other hand, are designated as the time point of minimal articulator closure (i.e. the time when the articulator starts to move towards the target). In Fig. 1, the kinematic data of the sequence *pak pak* are presented. Two major articulators are involved: Tongue Dorsum (TD; for /k/) and Lower Lip (LL; for /p/) (another being Tongue Tip (TT; for /t/)). The dashed line marks a gesture offset, and the solid arrow a gesture onset. An active gesture, represented with a shaded bar below, is defined as the span of gesture onset to gesture offset.

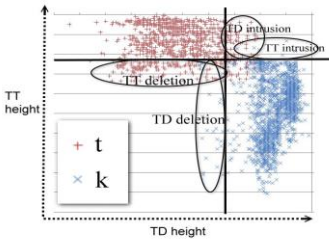
Figure 1: Defining gesture onsets and offsets.



2.4. Quantifying speech errors

Speech errors may refer to i) reduction error: an intended gesture is absent or reduced, ii) intrusion error: an unintended gesture shows up, and iii) substitution error: reduction and intrusion errors co-occur [2]. The error metric proposed in [3] is adopted here. That is, the interquartile range of, for example, TD height across all stop consonant trials *in either onset or coda position* for a participant was first calculated, and then the midpoint between two interquartile ranges (/k/ trials and non-/k/ trials) serves as the error threshold of the TD gesture. As we can see in Fig. 2, whereby all stop consonant trials in onset positions of Speaker 2 were plotted, if an intended /t/ (= a plus sign '+') is located in the first quadrant (i.e. the upper-right quadrant in Fig. 2), with TD height greater than the error threshold of TD, then that is taken as a TD intrusion on /t/. All that are located in the third (lower-left) quadrant are tokens with either no TD gesture no TT gesture; i.e., TD deletion if an intended /k/, TT deletion if supposed to be a /t/. Finally, the normal /t/'s and /k/'s, are in the second and fourth quadrants as well as substitution errors (i.e. a combination of full intrusion of unintended gesture and full reduction of intended gesture).

Figure 2: Scatterplot for TD and TT (in onset) for Spkr 2.

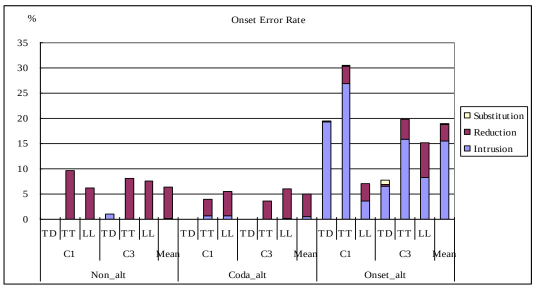


3. RESULTS

3.1. Errors in onset positions

Fig. 3 illustrates error rates in onset positions from the pooled data of four speakers. As shown, there are hardly any errors found in onset positions for the non-alternating (Non_alt; e.g. *pak pak*) and coda-alternating groups (Coda_alt; e.g. *pak pat*). There are some TT and LL reduction errors in Non_alt and Coda_alt, and these errors are mainly from Speaker 2 (see also S2 in Fig. 4).

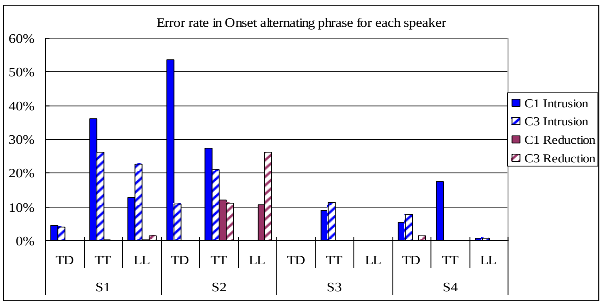
Figure 3: Error rates by error type across speakers in different types of disyllabic phrases (C1VC2 C3VC4).



Errors are more frequently found for the onset-alternating phrases (Onset_alt; e.g. *kap tap*) and notably, (TT and TD) intrusion errors are the dominant error type (confirming [2, 3]), while deletion and substitution errors are so rare that they could be ignored. TD and TT intrusion errors are more frequent in C1, whereas there are more LL intrusion errors only in C3.

In Fig. 4, it is obvious that there seem to be no consistent error patterns across speakers, if we compare the most “error-free” Speaker 3 and the most “errorful” Speaker 2. But, two tendencies can be noticed: i) TT intrusions are more likely to occur [2], and ii) a general preference may be observed for anticipatory intrusion in C1 to progressive intrusion in C3.

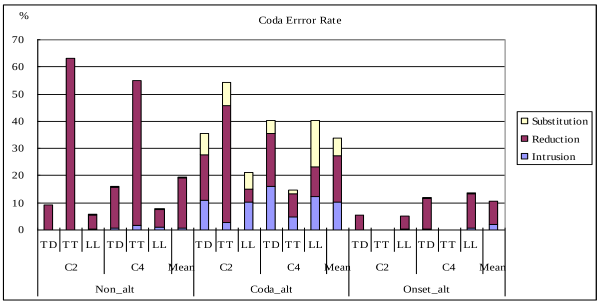
Figure 4: Error rates by error types for each speaker in onset-alternating phrases (C1VC2 C3VC4).



3.2. Errors in coda positions

Fig. 5 illustrates error rates in coda positions from the pooled data of four speakers.

Figure 5: Error rates by error types across speakers in different types of phrases (C1VC2 C3VC4).



It is expected that there would be fewer errors in coda positions for the non-alternating group, where C2 and C4 are identical (Non_alt, e.g. *pat pat*). In Fig. 5, however, our results surprisingly show that there are a fair amount of errors, suggesting that non-alternating conditions may not be always “error-free”. Notably, errors found under the current condition are overwhelmingly reduction errors, with very few instances of intrusion and substitution errors (whose error rates are less than 1%).

Furthermore, a general pattern can be observed with respect to the reduction rates of the gestures in Table 1: TT > TD > LL (where > means ‘more than’), while what all the four speakers have in common is that coda /p/’s are most resistant to reduction (except for Speaker 2). Finally, it should be noted that Speaker 2’s error rates of TT reduction reached nearly 100%, indicating that there were more “non-errors” among “non-control” stimuli for this particular speaker. It remains to be seen why that would be the case.

Table 1: Reduction errors in non-alternating phrases (C1VC2 C3VC4) for all speakers. (%)

Non-Alt	S1		S2		S3		S4	
Reduction	C2	C4	C2	C4	C2	C4	C2	C4
TD	2	40.4	6.4	0.8	1	0	28.2	12.7
TT	52.6	48.7	98.6	98.6	28.1	18.8	45.7	2.9
LL	1.3	4	14.3	15	0	1.6	0	0

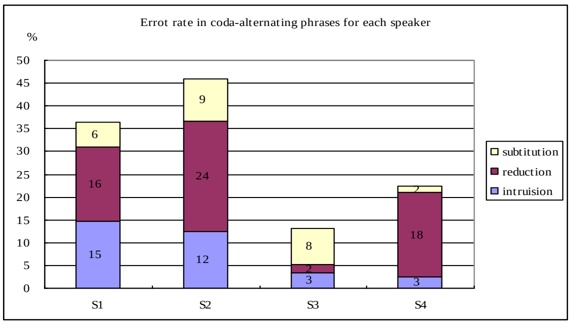
In contrast, in the onset-alternating phrases (e.g. stimuli like *kap tap*), the non-alternating codas (e.g. /p/’s in *kap tap*) are resistant to reduction (except for Speaker 4). Again, we see TD > LL in Table 2 (except for Speaker 2, again), a similar pattern we have seen in Table 1.

Table 2: Distribution of reduction errors on non-alternating codas for onset-alternating trials. (**%**) (Note that TT data are not available.)

Onset-alternating	S1		S2		S3		S4	
Reduction	C2	C4	C2	C4	C2	C4	C2	C4
TD	1.1	28	3.2	1.7	4.2	5.8	17.3	4.2
LL	0	2.4	13.2	29.6	0.8	3.1	1.4	5.8
Mean	0.6	15	8.2	15.7	2.5	4.4	9.4	5

Finally, errors are more frequently found in the coda alternating phrases (e.g. *pak pat*), as expected.

Figure 6: Error rate by error type for each speaker in coda-alternating phrases.



Remarkably, more reduction errors can be found, rather than intrusion ones in Fig. 6 (cf. [3]; except for Speaker 3). Also, a recurrent pattern is found in Table 3 below. That is, the reduction errors of coda /t/’s and coda /k/’s outnumber that of coda /p/’s to a great extent, i.e. TT > TD > LL, except for Speakers 1 and 2 (TD > TT in C4), if compared with errors in coda positions among the non-alternating group (see also Table 1).

Finally, different tonal patterns (i.e. H-L or L-H) are not correlated with specific speech errors.

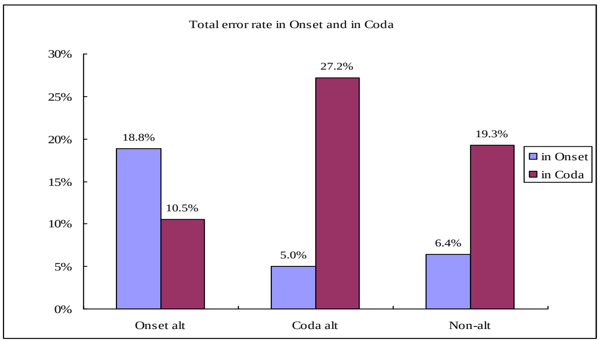
Table 3: Distribution of reduction errors on non-alternating codas for coda-alternating trials. (%)

Coda-alternating	S1		S2		S3		S4	
Reduction	C2	C4	C2	C4	C2	C4	C2	C4
TD	3.8	26.3	35.4	24.4	0	1	19.3	14.1
TT	61.9	15	53.9	5.8	7.6	0	46.2	9.4
LL	1.4	1.7	10	24.6	0	0	0	0

4. DISCUSSION

Our results of the Taiwanese data have shown that the bias of intrusion errors is robustly attested in onset positions (confirming [2, 5]). Furthermore, trials with alternating codas are more error prone than those with alternating onsets (confirming [3]).

Figure 7: Error rates in onset and coda positions.



A glaring cross-linguistic difference, however, is evidenced in coda positions. As we have learned, reduction errors are robustly attested in coda positions in both non-alternating and alternating conditions (e.g. *pak pak* and *pak pat*, respectively), while intrusion errors slightly outnumber reduction errors under the same circumstances in English [3]. Also, this finding is surprising in that it has been confirmed in English that non-alternating phrases are immune to errors as gestures are locked in 1:1 (stable) coordination mode [2, 5]. Moreover, as we have learned from Tables 1~3, it is evident that reduction errors occur irrespective of whether gestures oscillate in 1:1 (stable) or 1:2 (unstable) modes.

In contrast, intrusion errors are affected by coordination modes of gestures. More precisely, intrusion errors are found mainly on tokens with alternating consonants (in onset or coda positions; [2, 3, 5]), lending support to the view according to which unstable coordination between gestures (e.g. 1:2) is one of the sources of speech errors [2, 3]. In particular, the TT intrusion bias was found as the major error type in Taiwanese, too.

We believe that our principal finding (or the "onset-coda" asymmetry in speech errors) does not really challenge the gestural view of errors [2]; instead, the reduction bias found in Taiwanese can be attributed to this factor: as we have mentioned at the outset, stop codas in (South) East Asian languages are never released. We conjecture that, in checked syllables, the unreleasedness has a strong bearing on canonical/proper implementation of gestures because maximal constrictions may not be necessarily formed. In other words, gestures tend to be reduced in coda position (and notice that stop codas are *never* debuccalized since /ʔ/ is a possible coda), as maximal constrictions are *not* required (in order to produce a release burst). This may also be the situation with /t/-final word pairs (e.g. *cop cot*) in English dialects with glottalized allophones. That being the case, it may not be unexpected to see that reduction errors prevail in coda position. Last but not least, the effects of the accelerating production rate on the observed error count should also be factored in.

In conclusion, our results do not only provide support for the gestural view of speech errors in an East Asian language (at least, with regard to errors in the mismatching onset condition), but also enrich the typology of speech errors. To recap, the observed reduction bias is attributable to the fact that stop codas are implemented differently in between English and Taiwanese.

5. ACKNOWLEDGEMENTS

This study is supported by Grant DC 008780 to Louis Goldstein from NIH, NIDCD, USA and Grant NSC 98-2410-H-007-042 from NSC of Taiwan.

6. REFERENCES

[1] Edmondson, J., Chang, Y., Huang, H., Hsieh, F., Peng, Y. 2010. *Voiceless Applosive Finals in Taiwanese and Other E&SE Asian Languages: Laryngoscopic Case Studies*. Paper presented at *LabPhon 12* Albuquerque, NM.

[2] Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., Byrd. D. 2007. Dynamic action units in slip in speech production errors. *Cognition* 103(3), 386-412.

[3] Mooshammer, C., Goldstein, L., Tiede, M., Nam, H. Kulshreshtha, M., McClure, S., Dhillon, R. 2010. *Effects of Competition on Speech Production*. Paper presented at the *Conference on Motor Speech* Savannah, GA.

[4] Perkell, J., Cohn, M., Svirsky, M., Matthies, M., Garabieta, I., Jackson, M. 1992. Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *JASA* 92, 3078-3096.

[5] Pouplier, M. 2008. The role of a coda consonant as error trigger in repetition tasks. *JP* 36, 114-140.