ARTICULATORY SPEECH ERRORS AND WORD STRUCTURE

Anneke Slis, Rohan Bali, Aravind Namasivayam, Dimitra Chaldi, Pascal Van Lieshout

University of Toronto
anneke.slis@alum.utoronto.ca

ABSTRACT

This study explores how identical final CV syllables in C1VC3V C2VC3V word pairs affect speech error patterns. Recent kinematic studies have shown that speakers frequently added non-intended movements to intended articulatory movements during the first consonants of C1VC3 C2VC3 syllables or reduce the range of the intended movements [1, 2, 3]. This effect was especially strong in syllables that contained identical final consonants, such as in switching from "cop top" to "top cop" [2]. It was hypothesized for the current study that C1VC3V C2VC3V word pairs, in which the second and fourth syllable were identical, would show similar intrusion and reduction patterns. Articulatory movements of the tongue tip, dorsum and lower lip were recorded with the 3D Electro-Magnetic Articulograph (AG501). Stimuli consisted of pairs of open and pairs of closed syllables. The findings revealed that CVCV CVCV word pairs resulted in significant more intrusions than CVC CVC word pairs.

Keywords: kinematic data, intrusions and reductions, open and closed syllables, entrainment.

1. INTRODUCTION

For most of us, speaking seems an easy task which is performed more or less flawlessly. However, once in a while things go wrong and we make an error, like saying "cop top" where the speaker intended to say "top cop". When we *listen* to errors like this, the onset "c" of "cop" seems to have moved to the word "top", replacing the "t". Instead of examining errors perceptually, the line of research in our project builds on findings from recent studies which measure the actual articulatory movements of the tongue tip, tongue body, tongue dorsum, and the lips [1, 2, 3]. In these studies, speakers repeated syllable pairs with alternating onset consonants, such as "cop top", around 15 times. It was revealed that many errors did not conform to correct speech segments. Instead, movement errors were measured which seemed to result from a blend of two simultaneously produced speech segments [1, 2, 3]- In the case of "cop top", an extra tongue tip movement was added during the /k/ and an extra tongue dorsum movement was added during the /t/. Another crucial finding was that the number of errors increased during the 15 repetitions of a syllable pair. Especially closed syllable pairs, such as "cop top", showed this behaviour. Open syllable pairs such as "ka ta" or "ta ka ta", are found to be less error prone [2].

These error patterns resemble patterns of behaviour, observed in studies on motor control such as in finger tapping and drumming [see e.g., 4, 5]. Tapping a rhythm of 1 tap with one hand against 2 taps with the other hand is more difficult than a 1 against 1 rhythm [4]. The participants switch to simpler tapping-rhythms after many repetitions under faster rate conditions. This phenomenon is known as entrainment. The tapping studies showed that the higher frequency ratio acts as the "attractor": a 2:1 rhythm changes to 2:2 [5]. Extending this mechanism to speech implies that moving a tongue tip once for /t/ in "top", and moving the lips twice for /p/ in "top cop" is more difficult than moving both the tongue tip and lips twice, such as in "top top". This results in adding a tongue tip movement during /k/ in the first syllable for "cop top".

Entrainment is always present in moving systems, regardless of the properties of the individual components, such as for example the mass of the individual articulators. However, studies have shown that different combinations of consonants and vowels can result in different intrusion patterns [3]. In addition, open syllable pairs such as "ka ta" or "ta ka ta", are found to be less error prone [2].

One of the properties in which languages differ is the frequency of closed and open syllables. The Greek language consists mainly of open syllables, such as "ka" and "ta"; English is characterized by many closed syllables, like "cop" and "top". Because of these differences in syllable structure, it can be hypothesized that error patterns differ in languages that contain mainly open syllables compared with languages which contain more closed syllables. The data which are presented in this paper are part of a larger study investigating

differences between Greek and Canadian English coordination patterns.

This paper reports data collected from Greek speakers who produced CVCV CVCV and CVC CVC word pairs. CVCV CVCV word pairs with identical second and fourth syllables, such as for example “topo kopo”, have the same rhythmic characteristics as “cop top”. In this example, the “po” is the syllable that is produced twice, compared to “to” and “ko”. Thus a 1:2 relation exists between the first syllable (“to” and “ko”) versus the second syllable (“po”). This raises the question whether syllable strings that are characterized by different first open syllables but share a second open syllable, like for example “topo kopo”, result in similar intrusion and reduction patterns as words like “cop top”.

2. METHOD

2.1. Participants

3 male monolingual English and 3 male monolingual Greek speakers, without speech, language, or hearing impairments, were recruited. Results for a set of word pairs produced by three Greek speakers will be reported. Collection of speech from Canadian English Speakers is in progress.

2.2. Stimuli

One set of stimuli contained open syllable pairs, such as “kopo topo” (CVCV CVCV). The second CV in the first and second word were identical. A second set consisted of closed syllable pairs, such as “cop top” (CVC).

2.3. Procedures

A repetitive speech task was employed in which the participant was instructed to repeat the syllable pairs as often as possible; no limitation on the number of repetitions was set. Because speaking rate influences the number of intrusions and reductions [1, 3], two speaking rates were employed: normal (N) and fast (F). In one condition, the rate was controlled by a metronome (M+). In a second condition, the words were produced without a metronome (M-). For both the M+ and M- condition, the trial was preceded by 8 beats to indicate the correct speaking rate for the participant.

Kinematic data were collected with the Electromagnetic Articulograph (AG501). Movements of the tongue tip, tongue dorsum, lower lip and jaw were recorded at 250 Hz. At the same time, the acoustic signal was recorded (48 kHz) with a microphone built into the EMA. The collected movement data were processed with a Matlab toolbox developed at the Oral Dynamics Lab at the University of Toronto.

2.4. Error definition

The string of repetitions of the syllable pairs was segmented into separate syllable frames based on the minima of the coda-consonant. In case of the CVCV CVCV word pairs, the minima of the consonant in the second syllable were taken as the point of segmentation (in the example “topo kopo” this is the minimum of the bilabial closure /p/). From each frame the maximum displacement value of the target articulatory movement was calculated. This maximum target articulator value, e.g., the tongue dorsum during the onset of the syllable “cop”, served as the location at which the displacement of the non-target articulator was determined. In the word “cop”, for example, the non-target articulator is the tongue tip, which forms the target articulator in the word “top”.

To identify reductions, medians and MAD’s were calculated based on all the maximum displacement values of target articulatory movements within a trial (for 15 repetitions of the syllable pair “cop top” 13 values were used because the first and last repetition were always disregarded). Intrusions were calculated based on the values retrieved for the non-target articulator. Two MAD’s or higher than the median non-target value was labelled an intrusion; two MAD’S or lower than the median target value was labelled a reduction [3].

2.5. Statistics

The ratio of intrusions and reductions was calculated to correct for the fact that speakers produced trials of with a different number of repetitions (i.e., #intrusions/#repetitions). A sign test was employed in which the ratio of intrusions on “topo kopo” for each participant in each condition (with and without metronome in fast and normal speaking rate) was compared with the ratio of intrusions in “cop top”. An additional sign test was performed for the ratio of reductions. The ratio of intrusions and reductions were collapsed across position within the word pair (first or second word).

3. RESULTS

The sign test (3 participants, 4 conditions) showed that, compared to "cop top", 10 "topo kopo" words resulted in a higher ratio of intrusions, and 2 in a lower ratio of intrusions. This resulted in a significant difference in ratio of intrusions between "topo kopo" and "cop top" at the $p < 0.05$ level: $p <= 0.0386$ (see figure 1). For reductions the difference between "topo kopo" and "cop top" turned out not to be significant at the $p < 0.05$ level (n+ 8, n- 4): $p <= 0.388$ (see figure 2).

Figure 1: Ratio of intrusions is displayed on the vertical axis; G3, G5 and G6 are the three Greek speakers. FM+ is the fast condition with metronome; FM-: fast condition without metronome; NM-: normal speaking rate without metronome; NM+: normal speaking rate with metronome.

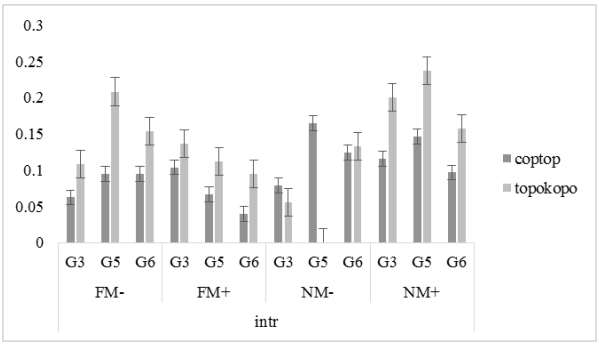
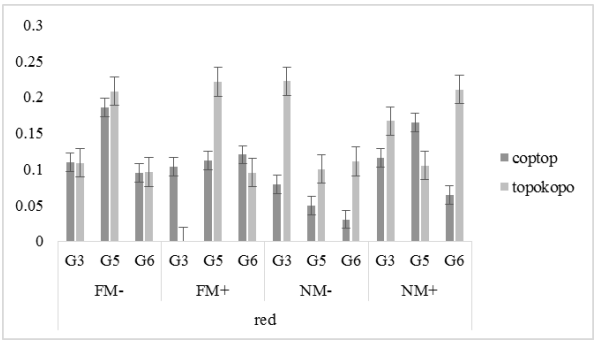


Figure 2: Ratio of reductions is displayed on the vertical axis; G3, G5 and G6 are the three Greek speakers. FM+ is the fast condition with metronome; FM-: fast condition without metronome; NM-: normal speaking rate without metronome; NM+: normal speaking rate with metronome.



4. DISCUSSION

The pilot study showed that speakers produced more intrusions on the word pair "topo kopo" than "cop top". It can be argued that differences in ratio of intrusions between "topo kopo" and "cop top" resulted from the fact that the first word pair is an existing word in Greek, whereas "cop top" is a nonword. However, familiarity with the words should result in a more stable coordination pattern and thus fewer intrusions. This analysis showed the opposite pattern. Comparison with data from Canadian English speakers can shed light on the role of familiarity on the occurrence of intrusions and reductions. Collection and analysis of data from Canadian English speakers is currently in progress.

It can be inferred from the preliminary data that the second and fourth syllable in "topo kopo" acts as the higher frequency component comparable to the coda consonant in "cop top".

5. REFERENCES

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

[2] Pouplier, M. (2008). The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics, 36*, 114-140.

[3] Slis, A. W. (2014). *The Effect of Articulatory Constraints and Auditory Information on Patterns of Intrusions and Reductions.* (Unpublished doctoral dissertation), University of Toronto, Canada.

[4] Peper, C., & Beek, P. (1998). Distinguishing between the effects of frequnecy and amplitue on interlimb coupling in tapping a 2:3 polyrhythm. *Experimental Brain Research, 118*, 78-92.

[5] Peper, C., Beek, P., & Van Wieringen, P. (1995). Multifrequency coordination in bimanual tapping: Asymmetrical coupling and signs of supercriticality. *Journal of Experimental Psychology, 21*(5), 1117-1138.