Effects of phonological competition on speech planning and execution

Christine Mooshammer[1,2], Mark Tiede[2], Argyro Katsika[2], Louis Goldstein[2, 3]

[1]IdSL HU Berlin, Germany [2]Haskins Laboratories, New Haven, U.S.A., 3Department of Linguistics, USC, Los Angeles, U.S.A.
christine.mooshammer@hu-berlin.de, tiede@haskins.yale.edu, katsika@haskins.yale.edu, louisgol@usc.edu

ABSTRACT

Competition between phonologically similar sequences in an utterance is one of the major causes for speech production errors. Additionally, phonological competition has been found to increase planning time and slow down speech rate in CVC word pairs. The aims of this study are to investigate the timecourse of phonological competition by employing different tasks and by a detailed gestural analysis. Effects of competition in the onset ("top cop" and "pay Kay") are compared to competition in the coda (e.g. "top tock") and both were compared to sequences of identical words (e.g. "top top"). Results from three studies are reported: acoustic latencies from a delayed naming task and a simple naming task (18 speakers), and articulatory latencies from a delayed naming task using EMA (6 speakers). Reaction time was affected by competition but not by locus. Mismatch in the coda lengthened the execution time most prominently in the final rime.
Keywords: speech planning, phonological competition, gestural timing, syllable structure.

1. INTRODUCTION

It has long been established that similar sequences are more difficult to say than dissimilar ones, shown by longer planning and execution times [7, 10, 12] and by a higher error rate [e.g. 11]. Similarity in this context refers to two characteristics, the segmental context and the number of shared features. For example, an error-eliciting environment consists of sequences with identical phonemes but also with phonemes that are only similar and not identical, as exemplified in tongue twisters such as *She sells sea shells*. Traditionally, higher error rates and longer planning and execution times have been attributed to competition between simultaneously activated phonemes due to spreading activation [1]. Errors occur when phonemes get mis-selected because they are activated at the same time on the phonological level: the highest-activated phoneme will be articulated, and competing ones not. This leads to categorical phoneme sized errors on the phonetic output level. Within the more recent Cascading Activation model the activation of all co-active phonemes cascades to the motor level and therefore it can account for the occurrence of gradient, non-canonical errors [3].

Both accounts assume some kind of baseline co-activation by repeating parts of words. But this does not explain asymmetries due to syllable position: Sevald and Dell attributed more frequent coda errors to re-activation within the sequential cuing model [SCM, 10]. For word sequences such as *top tock* with a mismatch in the coda, saying the second word *tock* reactivates the first word *top* because of their shared initial segment sequence. When the speaker reaches the /k/ in *tock*, the reactivated coda /p/ of *top* competes with the /k/ and an error might occur. This does not happen for the onset mismatch such as *top cop* because *cop* does not reactivate *top*. This mismatch asymmetry affects execution time and error rates but probably not planning time.

The aims of this study are to investigate effects of similarity and position within the syllable in several non-repetitive tasks. The first experiment consisted of a simple naming task comparing CV and CVC word repetitions (e.g. *pay pay*, *top top*, hereafter: EQ) to words with different stops in the onset (*pay Kay*, *top cop,* hereafter: OD) to words with different codas (CVC only, *top tock* CD). In order to exclude effects of lexical variables Experiment 2 was a delayed naming task with the same stimuli as in Experiment 1. Since the majority of final stops were not released execution time could not be measured accurately by means of audio data. Therefore, in Experiment 3 movement data were acquired for the same task as in Experiment 2.

2. EXPERIMENT 1: SIMPLE NAMING TASK

2.1. Method

Participants: Eleven male and seven female native speakers of American English participated.
Stimuli: The material, shown in Table 1, consisted of 6 sets of CV and 8 sets of CVC words pairs differing in mismatch in the onset (OD) and for the CVC also in the coda (CD). Identical pairs (EQ) served as controls. The initial consonants were voiceless stops, while final consonants could be either voiced or voiceless stops. The vowel was always the same within a word pair.

The following lexical variables were obtained per word, based on the Celex Corpus: word and syllable frequency [6], phone and biphone probability [13].

Table 1: Examples of the material for experiment 1. EQ: identical words, OD: onset different, CD: coda different. Non-words are in italics.

CVC	EQ	OD	CD
	tape tape	tape cape	tape take
	top top	top cop	top tock
	pick pick	pick tick	pick pit
	cod cod	cod pod	cod cob
CV	EQ	OD1	OD2
	pay pay	pay *tay*	pay Kay
	ta ta	*ta* pa	*ta Kaa*

Procedure: In order to elicit fixed starting positions the participants were asked to produce a sustained schwa before saying the test words (e.g. [5]). Participants were shown a *Get ready, say 'uhhh'* prompt on a computer screen. Following a randomized delay between 1000 and 2000 ms the screen border changed to green, the test words appeared and an audible beep was emitted as cues to say the target words as quickly as possible. The inter-trial interval was 250 ms.

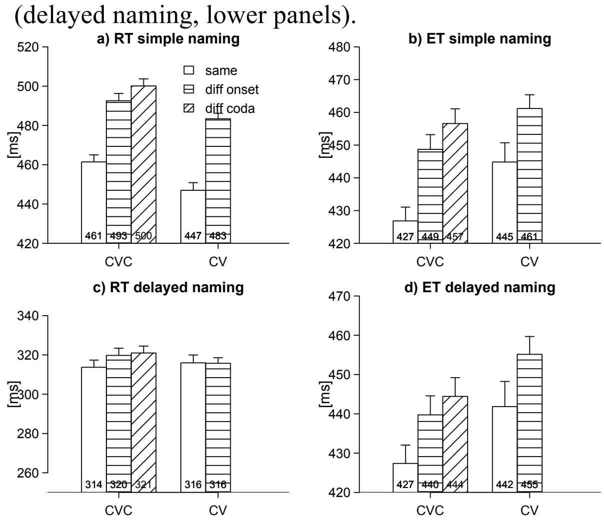
Measurements: The acoustically defined reaction time RT_{ac} was calculated as the interval from the peak of the beep to the release burst of the initial stop. The execution time ET_{ac} was measured as the interval from the release burst of the initial stop to the offset of the vowel of the second word. Since the final stop was frequently unreleased it could not be used as the offset of the word pair.

Statistics: Effects of similarity and mismatch location were tested by linear mixed effects models with subject and item as random factors, and similarity as a three-level factor for CVC (EQ, OD, CD) or two-level factor for CV (EQ, OD). The dependent variables were the log-transformed RT_{ac} and ET_{ac}. Effects of lexical frequency and phonotactic probability were tested by comparisons between models with and without these covariates.

2.2. Results

Fig. 1 shows the means and standard errors, estimated by using LME models, for RT_{ac} and ET_{ac} in the upper two panels. For both, CVCs and CVs, the RT_{ac} (Fig. 1a)) was significantly longer for similar words pairs compared to identical words (CVC EQ-OD: $t=7.5$, $p<0.001$, EQ-CD: $t=9.2$, $p<0.001$; CV EQ-OD: $t=6.2$, $p<0.001$). The location of dissimilarity (onset vs. coda) did not affect RT_{ac} significantly for the CVC pairs. Lexical variables affected the RT_{ac} significantly: For the CVC word pairs, word frequency ($t=-2.9$, $p<0.01$) and phone probability ($t=-4.3$, $p<0.01$) significantly sped up reaction times. For CV sequences word frequency had a significant facilitating effect ($t=4.2$, $p<0.001$).

Figure 1: Model-based means and standard errors of RT_{ac} (left panels) and ET_{ac} (right panels) for Experiment 1 (simple naming, upper panels) and Experiment 2 (delayed naming, lower panels).



Similarity increased ET_{ac} for dissimilar pairs (CVC EQ-OD: $t=5.6$, $p<0.001$, EQ-CD: $t=7.0$, $p<0.001$, CV EQ-OD: $t=4.0$, $p<0.001$). Furthermore, ET_{ac} was significantly longer for CD (e.g. *top tock*) than for OD (e.g. *top cop*). Higher syllable and word frequencies reduced the ET_{ac} significantly (CVC syllable frequency: $t=-4.2$, $p<0.001$, word frequency: n.s.; CV syllable frequency: $t=-2.4$, $p<0.05$, word frequency: $t=-3.8$, $p<0.001$).

In summary, for the simple naming task dissimilarity increased RT_{ac} and ET_{ac}, for both CV and CVC word pairs. The effect of location was only significant for ET_{ac} with longer execution times for mismatch in the coda as compared to mismatch in the onset. This result confirms the Sequential Cuing Model (SCM) [10] since competition due to coda mismatch should be more detrimental than onset mismatch and it should impede execution more than planning. However, both RT_{ac} and ET_{ac} were significantly influenced by several lexical variables. Since the phonetic composition of stimuli was constrained, these variables cannot be controlled in this experiment. Therefore, Experiment 2 employed a delayed naming task in order to exclude possible effects of lexical access [e.g. 8, 3]

3. EXPERIMENT 2: DELAYED NAMING TASK

3.1. Method

Procedure: Instead of simultaneous presentation of the stimulus and the audible beep signal as in the simple naming task, in the delayed naming task the stimulus was shown first and then after a randomized delay varying between 1000 and 2000 ms the audible beep was emitted. Participants, measurements, material and statistics were the same as in Experiment 1.

3.2. Results

The RT_{ac} was much shorter for delayed naming (Fig. 1c)) than for simple naming (Fig 1a)), confirming results from [8, 3]. Furthermore, the difference between conditions was much smaller (30 ms simple vs. 7 ms delayed naming) and only reached significance for the CVC pairs with longer delays for CD than EQ ($t=2.5$, $p<0.05$). There was no difference between RT_{ac} for identical and dissimilar CV pairs. As was expected none of the lexical variables affected the RT_{ac} in the delayed naming task.

The execution time ET_{ac} was similar to the simple naming task and significantly affected by word frequency ($t=-3$, $p<0.01$) indicating that more frequent word pairs were produced faster. Identical CVC word pairs were produced significantly faster than word pairs with mismatch in the onset (OD, $t=2.9$, $p<0.01$) and mismatch in the coda (CD, $t=4.6$, $p<0.001$). The difference between OD and CD was not significant (see Fig. 1d)).

In summary, differences in RT_{ac} were reduced in the delayed naming task due to the exclusion of effects of the lexical access. The execution time was affected in a similar manner in the two experiments with the longest ET_{ac} for CD, OD in-between and shortest for EQ. The difference between OD and CD is predicted by the SCM but it did not reach significance in Exp. 1 and 2. According to the SCM the lengthening effect of competition should be located mainly after the onset of the second word is executed, i.e. the rime. However, since about half of the final stops in the CVC word pairs were unreleased, only the effect up to the vowel offset in the second word could be measured. In order to access the full extent of lengthening and to get a more detailed picture of its temporal structure an EMA experiment with physiological speech data was carried out.

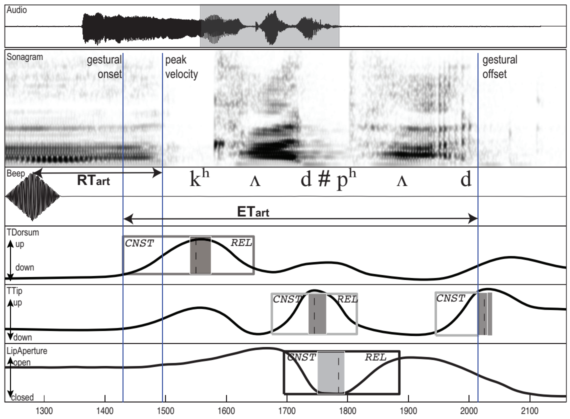
4. EXPERIMENT 3: DELAYED NAMING TASK WITH EMA

4.1. Method

Participants: Seven native speakers (4f, 3m) of American English were recorded by means of EMA with a delayed naming task and the same material as in Exp 2.

Apparatus: Tongue, jaw, lips and head movements were recorded by means of 3D Electromagnetic Articulography (AG500, Carstens AG, see [4]) with a 200 Hz sampling rate. The sensor trajectories were corrected for head movement and rotated to the occlusal plane.

Figure 2: Labeling of the articulatory data, shown for the word pair *cod pod*. Panels: audio signal and spectrogram, audio signal of the beep, vertical tongue dorsum and tongue tip movement, lip aperture. Measurements are explained in the text.

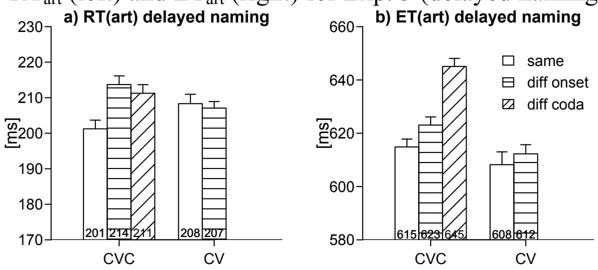


Measurements: Consonantal gestures (2 for CV word pairs, 4 for CVC word pairs) were labeled to identify a *constriction* interval (movement towards constriction, see boxes denoted with *CNST* in Fig. 2), a *constricted* interval (relatively stable closed phase, filled grey boxes in Fig. 2), and a *release* interval (opening movement, box denoted with *REL*) using velocity criteria applied to the trajectory of the most relevant sensor. For bilabial stops the Euclidean Distance between the sensors of the upper and lower lips was calculated. The reaction time based on articulatory data, RT_{art}, corresponds to the interval between the peak of the beep and the velocity peak of the most relevant sensor for initial consonant. The articulatorily defined execution time, ET_{art}, corresponds to the interval between the gestural onset of the initial consonant to the offset of the constriction interval of the final consonant gesture for CVC pairs and the offset of acoustic vowel for the CV pairs.

4.2. Results

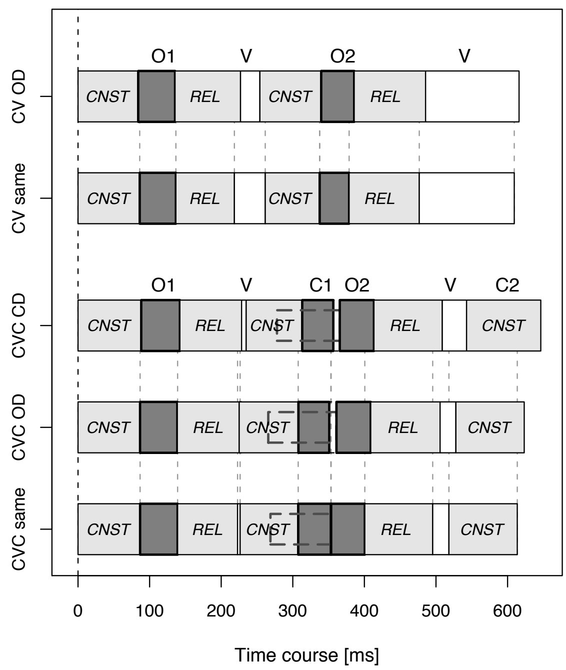
Figure 3 shows the model-based estimates of articulatory measurements RT_{art} (left) and ET_{art} (right). Similarity only had a significant effect on RT_{art} for CVC (EQ-OD: $t=3.0$, $p<0.01$, EQ-CD: $t=2.8$, $p<0.01$) but not for CV. For the CVCs items with CD were produced with significantly longer ET_{art} than EQ ($t=8.0$, $p<0.001$) and OD ($t=5$, $p<0.001$). The difference between OD and EQ is much smaller but still significant ($t=2.3$, $p<0.05$). For CV no lengthening due to mismatch in the onset could be detected. None of the lexical variables had a significant effect on RT_{art} or ET_{art}.

Figure 3: Model-based means and standard errors of RT_{art} (left) and ET_{art} (right) for Exp. 3 (delayed naming).



The SCM predicts that lengthening for CD should take place during the final rime or coda since it is triggered by reactivating the first coda after producing the onset of the second word. To test this, the gestural timing for the two (CV) or four consonants (CVC) is displayed in Fig. 4.

Figure 4: Gestural score for the CV sequences (upper two bars) and the CVC sequences (lower 3 bars). The light gray bars correspond to constriction (*CNTS*) or release movements (*REL*), the dark gray bars to the constricted phases. The white patches indicate non-overlapping parts, e.g. during vowels or between words. The box with dashed line is the constriction gesture of the initial stops of the second word.



For the CVCs the most prominent location of lengthening for the CD condition is indeed the rime (18.5 ms longer than EQ, $p<0.01$), as predicted by the SCM. Non-identical words showed significantly less overlap between the coda of the first word and the onset of the second word (see white patches between C1 and O2 for CVC OD and CVC CD in Fig. 4, $p<0.001$). Even though mismatch in the onset lengthened the ET_{art} only to a small degree, for both CVC and CV a significant lengthening of *REL* in O1 and *CNST* of O2 was found. These lengthening effects did not fully contribute to the overall ET because there was also more overlap for the OD condition. For CV, the first V duration was shorter for OD than for EQ contexts. For CVCs the constriction gestures of the second onset (dashed box in Fig. 4) were more overlapped with the codas of the preceding word.

5. DISCUSSION

The aim of this study was to shed light into the time-course of phonological competition and the processing levels that are affected. As was expected simple repetition of the same word had a facilitating effect on the planning and the execution time which was much larger for the simple naming than for the delayed naming task. This differential effect on the RT suggests that processes at the stage of lexical access increase the planning time for the second word before the first word is initiated in the simple naming task. Phonological encoding and motor planning for the second word delays the initiation of the first word by around 8 ms (effect size in Exp. 2 and 3). The location of mismatch did not affect the RT but did affect ET. As predicted by the SCM, mismatch in the coda delayed the execution of the final rime and therefore lengthened the execution time. However, by means of detailed articulatory analyses we also found longer constriction gestures for the initial consonant of the second word for mismatch in the onset. Since competition should mainly be triggered by repeating the same onset (i.e. as in the coda mismatch condition) this was not expected. This lengthening is not contributing to the ET because it is overlapped by the preceding word, increasing the overlap between C1 and O2 in CVCs. For CVs the gestural lengthening of O1 and O2 shortened the phase in-between (termed V). This might explain why the ET_{art} increase for onset mismatch in CVs did not reach significance.

However, it cannot explain why the RTs in the delayed experiments were very similar for identical items and onset mismatch for CVs but not for CVC. One possible reason could be that in CV sequences only the vowel is shared whereas in CVCs it is the rime. A similar asymmetry has been found by Pouplier [9], with much smaller intrusive error rates in alternating CVs than in alternating CVCs.

6. REFERENCES

[1] Dell, S. G. 1984. Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10, 222-233.

[2] Goldinger, S., Tamiko, A., Abramson, M., & Jain, P. 1997. Open wide and say ''blah!'': Attentional dynamics of delayed naming. *Journal of Memory and Language* 37, 190–216.

[3] Goldrick, M., & Blumstein, S. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes* 21, 649-683.

[4] Hoole, P., Zierdt, A., & Geng, C. 2003. Beyond 2D in articulatory data acquisition and analysis. *Proc. 15th ICPhS* Barcelona, 265-268.

[5] Kawamoto, A., Lui, Q., Mura, K., & Sanchez, A. 2008. Articulatory preparation in the delayed naming task. *Journal of Memory and Language* 58, 347-365.

[6] Levelt, W., & Wheeldon, L. 1994. Do speakers have access to a mental syllabary? *Cognition* 50, 239-269.

[7] Meyer, E. D., & Gordon, P. 1985. Speech production: Motor programming of phonetic features. *Journal of Memory and Language* 24, 3-26.

[8] Monsell, S. (1990). Frequency effects in lexical tasks: Reply to Balota and Chumbley. *Journal of Experimental Psychology: General* 119, 335–339.

[9] Pouplier, M. 2008. The role of a coda consonant as error trigger in repetition tasks. Journal of Phonetics, 36(1), 114-140.

[10] Sevald, C., & Dell, S. G. 1994. The sequential cuing effect in speech production. Cognition, 53, 91-127.

[11] Shattuck-Hufnagel, S., & Klatt, H. D. 1979. The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18, 41-55.

[12] Sternberg, S., Knoll, R., Monsell, S., & Wright, C. 1988. Motor programs and hierarchical organization in the control of rapid speech. *Phonetica* 45, 175-197.

[13] Vitevitch, M., & Luce, P. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers* 36, 481–487.