Is Segmental Interference Position-dependent?

Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Harrison, William
Hepner, Christopher R.
Nozari, Nazbanou

2020

Is Segmental Interference Position-dependent?

William Harrison (wharris2@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15232 USA

Christopher R. Hepner (chepner@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15232 USA

Nazbanou Nozari (bnozari@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University,
5000 Forbes Avenue, Pittsburgh, PA 15232 USA

Abstract

The paper investigates the existence of position-independent segments in written and typed word production. In two experiments, we employed the segmental interference effect to first replicate past findings that naming a picture is more difficult in the context of another picture with which it shares segments in the same position (e.g., glow-flow) compared to an unrelated word (e.g., glow-cave). We then tested a new condition, in which the same target word is paired with an anagram of the original competitor (glow-wolf). Critically, the anagram shared the same number of segments with the target word, but never in the same position. Both experiments found robust interference for targets produced in the context of anagrams, with a magnitude comparable to the interference induced by the position-overlapping word. The results suggest that not only are position-independent segments represented in the production system, but they also play a critical role in activating segmentally related words and creating competition during word production.

Keywords: word production; segmental encoding; positional frame; segmental interference

Introduction

One of the most important questions in models of language production is the nature of the representations involved in the process of mapping meaning to sound. Major advances in the field resulted from the discovery of separate lexical and segmental layers, with distinct stages of processing (e.g., Garrett, 1975), and a very large body of literature, both on spoken and written production, has since focused on determining the nature of additional representations in the production system (e.g., Levelt et al., 1999). A special challenge in this regard is the nature of segments (phonemes in spoken and graphemes in written production) which mediate the mapping between lexical items and motor commands to articulate the word or write it down. The challenge stems from the fact that segments, unlike higher-level representations such as lemmas, lexemes, and semantic features, must be ordered and produced in the correct sequence even at the level of single-word production. This problem can be readily seen in words such as "pot" vs. "top", which have the same segments but differ in where those segments appear. This problem has led to the proposal of position-dependent segments in production (e.g., onset /p/ vs. coda /p/ as opposed to the generic /p/). The question of whether position-independent segments are still represented in the production system or not remains. This paper addresses this question.

Representation of segments in models of language production

The nature of segments in language first came under scrutiny after Lashley's seminal article criticizing chaining as a viable account of serial order in language (Lashley, 1951). The first instance of an alternative, *context-sensitive coding*, was proposed by Wickelgran (1969), who proposed that segments are further specified by the environment in which appear. For example, /p/ in "pot" is /${}_{-}\mathrm{p}_{ɒ}$/, while /p/ in "top" is /${}_{ɒ}\mathrm{P}_{-}$/. The two are thus distinct representations, distinguished by the attachments which represent their context within the word. Apart from requiring a large number of segments, this account was unable to explain findings like the strong tendency for segmental migration errors to maintain their positions within syllables: onsets are much more likely to replace other onsets than codas, and vice versa. This finding gave rise to the proposal of segmental "frames", i.e., abstract slots to which segments are bound. This kind of representation was implemented in models like Dell (1986) as onset, nucleus, and coda clusters for consonant-vowel-consonant (CVC) syllables such as "pot". In this method, /p/ would still have different representations in "pot" and "top" (as /$\mathrm{p}_{\mathrm{onset}}$/ and /$\mathrm{p}_{\mathrm{coda}}$/, respectively), but its advantage over Wickelgran representations was that /$\mathrm{p}_{\mathrm{onset}}$/ was the same for all words that started with /p/, and had the privileged status of replacing other segments that were also attached to the onset slot in the syllabic frame.

Since then, this method of frame-bound representations has been common in a variety of language models, although "frame" can represent different constructs, like syllabic structure or position in the word. Examples include models of speaking (e.g., Foygel & Dell, 2000), spelling (Houghton & Zorzi, 2003), and reading (Harm & Seidenberg, 1999 and its predecessors). This implementation is, of course, to some extent a matter of computational simplification. But if position-independent representations are essential for explaining fundamental aspects of word production, such as the dynamics of facilitation and interference due to the activation of competitors, then this simplification has non-negligible consequences. We approach this question of position-independent segmental representations in writing and typing by comparing interference between words that share segments in either the same or different positions.

Segmental interference in word production

For years, overlap in segments was thought to facilitate production. In addition to priming paradigms, blocked cyclic naming paradigms, in which a small set of pictures were to be repeatedly named, have been used to show that the same target was named faster if it shared its onset with other words in the cycle (e.g., pig, pen, pot) compared to when it did not (e.g., pig, bed, sun; e.g., O'Séaghdha & Frazer, 2014). This effect, however, turned out to be strategic, and turned into interference when overlap was moved to non-onset segments (e.g., mat/hat; Nozari et al., 2016), or when onset-overlapping words were interleaved with words that overlapped in non-onset positions (Breining et al., 2019). The interference was robust in both spoken and written production.

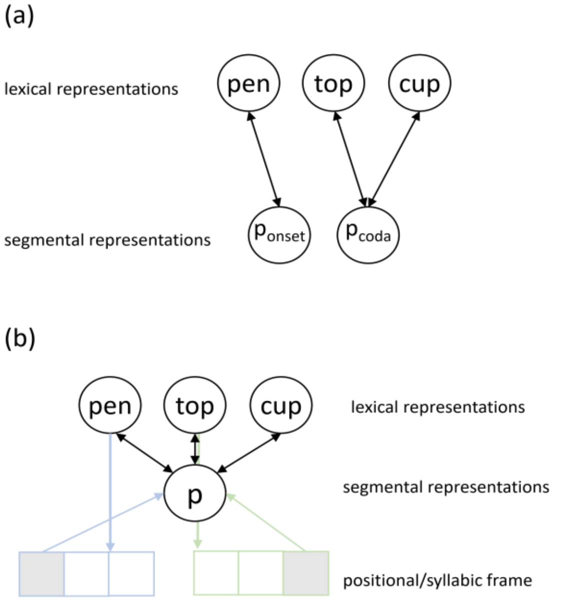

Figure 1: Schemata of the position-dependent (a) and position-independent (b) coding of segment /p/ in "pen", "top" and "cup". The blue and green boxes show the positional/syllabic frames for "pen" and "top", respectively. Specific slots activate segment /p/ based on its position in the word, e.g., through competitive queueing mechanisms (Houghton, 2018). Critically, model (b) has a position-independent representation of segment /p/ which activates both "pen" and "cup" when "top" is activated. No such representation is available in model (a), thus only "cup" is activated by "top".

The basic mechanism for segmental interference is that the segments activated by the target word feed back to the competing word, further increasing its activation and subsequently the activation of its phonemes, which compete with those of the target word. When the competing word shares no segments with the target word, the feedback does not activate it, so there is much less interference. This is true irrespective of whether segmental interference results from spreading activation or from incremental learning (Breining et al., 2019).

The interference induced by segmental overlap can be used to study the nature of the segmental representations involved in word production. So far, all evidence for segmental interference in picture naming comes from words that share segments in the same position. The resulting interference is thus compatible with position-dependent representations: /p_{coda}/ in "cup" activates /p_{coda}/ in "top", making it a stronger competitor than an unrelated word such as "bed" (Fig. 1a). The question is what happens if "top" is instead paired with "pen". If the representation of /p/ is strongly position-dependent, then the two /p/s in "pen" and "top" are different /p/s (/p_{onset}/ vs. /p_{coda}/) and should not cross-activate one another (Fig. 1a). Therefore, the interference induced by "pen" on "top" should be close to that of an unrelated word, and far less than that induced by "cup". If, on the other hand, a position-independent /p/ is still strongly represented in the system, then "pen" should interfere as strongly with "top" a5s would "cup" (Fig. 1b). We test these predictions in two experiments, one in handwriting and one in typing, pairing the same four-letter target word with an unrelated word (Baseline condition), a competitor that shares three letters in the same position (Overlap condition), and a second word that is the anagram of the competitor in the Overlap condition, i.e., it still shares three segments with the target, but in different positions (Anagram condition). Since segmental interference has been reported in response times (RT) and/or durations, we measure both in responses to the same target picture when it is paired with each of the three types of competitors. Based on prior findings, we expect interference in the Overlap vs. Baseline condition. An equally strong interference effect in the Anagram condition would point to a strong representation of position-independent segments in the production system. Absent (or much weaker) interference in the Anagram condition would rule out the strong representation of position-independent segments.

Experiment 1

Methods

Participants Forty-two native English-speaking undergraduate students (23 female, 18 male, 1 non-binary; mean age = 19.36 (SD = 1.32); 40 right-handed) participated for course credit.

Stimuli Stimuli consisted of 12 four-letter target words (e.g., *glow)* paired with a second four-letter word (henceforth referred to as "competitors") in three conditions (Overlap, Anagram, and Baseline). In the Overlap condition, the competitor shared all but the initial segment with the target (e.g., *flow*). In the Anagram condition, the segments of the word used in the overlap condition were rearranged to create an anagram (e.g., *wolf)*[1]. In the Baseline condition, target words were paired with words that did not share segments or CV structure with either the target or anagram (e.g., *cave*). Stimuli were balanced for frequency across condition using frequency data from the SUBTLEX-US corpus. Color images corresponding to the 48 words in the experiment were selected from Google images and sized to 320x320 pixels.

The 12 targets, each appearing with a competitor in three conditions, created 36 blocks. The order was counterbalanced such that the first, second, and third appearances of the targets were equally distributed across the three experimental conditions. With this constraint, six lists with pseudo-randomized block orders were created, with the same target never appearing in two adjacent blocks.

Procedure The experiment was developed in jsPsych (de Leeuw, 2015) and administered in an internet browser running on a PC and displayed on a Huion Kamvas Pro 12 tablet, on which participants also wrote their responses. Prior to beginning the experiment, participants saw and labeled four practice images, presented one at a time, to get comfortable using the tablet.

Participants were then assigned to one of the six lists and completed 36 blocks of word pairs as described above. At the start of each block, participants were shown a pair of images and their corresponding labels (e.g. *glow-flow*), and practiced the labels until comfortable, to reduce imageability effects. From this point on, only one image at a time was presented, and participants were instructed to write down the label as quickly and accurately as possible on the tablet. In each trial, a fixation cross was presented in the middle of the screen for 700 ms. The image was then presented in the center of the screen along with a 1×2.5-inch response box underneath. The image remained on the screen for 2000 ms, or until a response was initiated. Participants then had 2000 ms to complete their response. After that, the fixation cross for the next trial would immediately appear on the screen. At the beginning of each block, participants completed 4 practice trials with the pictures of that particular block. They then completed 16 experimental trials (8 presentations of each image). Trials in a block were pseudo-randomized such that an image did not appear more than twice in a row. Across all blocks, a total of 576 responses were collected in the experimental trials. The experiment took approximately 50 minutes to complete, including two short breaks.

Analyses Analyses were conducted in R (version 3.6.1; R Core Team, 2019). Trials with incorrect or null responses, RTs of less than 200 ms, or RTs more than 3 standard deviations away from each participant's mean were excluded from the analysis. To ensure the tightest experimental control, analyses focused on target items, which could be compared to themselves in different conditions[2]. Three planned contrasts on RTs and durations were tested: Overlap minus Baseline, Anagram minus Baseline, and Overlap minus Anagram. Data were analyzed using two methods with complementary strengths for ensuring the reliability of results. 1) For each contrast, we compared the absolute value of the mean difference in target RTs and durations across participants to the distribution generated by a Monte Carlo simulation with 100,000 permutations, resampling within participants such that order of appearance (1st, 2nd, or 3rd block containing the target) was maintained. This method is conservative and does not require assumptions about the shape of the distributions, etc., but does not capture the nested structure of items under subjects, or the potential item effects. 2) To ensure that the observed effects were not item-specific, the data were also analyzed with linear mixed effects models (LMEMs) using lme4 (version 1.1-21; Bates et al., 2015), in conjunction with lmerTest (version 3.1-0; Kuznetsova et al., 2017) to calculate p-values using Satterthwaite approximations. RTs were log-transformed to better approximate a normal distribution, and all numeric variables were centered and scaled. Comparisons of interest were contrast-coded in the models. The models included random intercepts for participants and targets. They did not tolerate random slopes.

Results & Discussion

Out of 12,096 total trials containing target items, 4.5% were excluded based on the criteria described above. Error rates were low and similar across conditions (0.9% of Baseline trials and 1.0% of both Overlap and Anagram trials). After excluding these trials, 11,546 trials were included in the analyses.

Figure 2 shows the RTs and durations for the target words in the three conditions in Exp 1. Analyses of targets revealed that RTs were not significantly longer in the Overlap vs. Baseline condition ($M = 1.28$ ms, $p = .632$). They were longer in the Anagram vs. Baseline condition ($M = 5.43$ ms, $p = .043$) but the difference between Overlap and Anagram was not significant ($M = -4.14$ ms, $p = .125$). Durations, on the other hand, were significantly longer in both Overlap and Anagram conditions ($M = 8.67$ ms, $p = .048$, and $M = 15.02$ ms, $p < .001$, respectively) compared to Baseline, but again the difference between Overlap and Anagram conditions was not significant ($M = -6.34$ ms, $p = .146$). The results of the LMEMs were identical, except that the non-significant difference between durations in the Overlap and Anagram conditions became significant in this analysis ($p = .036$).

To summarize, we found evidence of interference in handwriting durations in the Anagram condition compared to

---

[1] Because of the limited number of 4-letter words with anagrams, imageability could not be matched at the trial level, but was balanced in the experiment.

[2] LMEMs on a dataset that included both targets and competitors returned similar results to those with only target items.

the Baseline, which, despite its small effect size, was highly robust in both methods of analyses. Moreover, both methods showed significantly longer RTs in Anagram compared to Baseline conditions. Together, these results point to a clear interference effect induced by words that share segments with the target, even when there is no positional overlap between any of the segments.

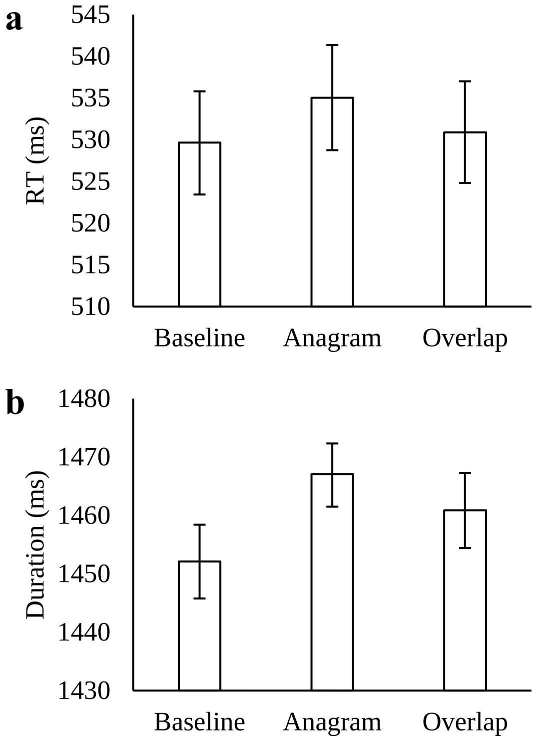

Figure 2: Mean (a) RTs and (b) durations for responses to target items in Baseline, Anagram, and Overlap conditions across participants in Experiment 1. Error bars represent 95% CIs, not corrected for between-subject variance.

Several findings, however, called for a replication and further investigation of the effect. First, the effect of interference on RTs was absent in the Overlap condition (cf. Nozari et al., 2016). The reason could be that handwriting is not ideal for exploring RTs, because participants could put the pen on the pad (and register an RT) before having decided exactly what letter to write. Typing, on the other hand, does not have this problem, since the identity of the letter must be determined before pressing the first key. Second, the comparison of the effect between Anagram and Overlap on durations yielded inconsistent results across the two methods of analysis. LMEMs found significantly longer durations in the Anagram condition, but Monte Carlo simulations did not. Although the direction of the effect, which, if anything suggests greater interference in the Anagram compared to Overlap conditions, does not call the main conclusion into question, it is worthwhile establishing whether it is a true effect. If it is, we hypothesized that this difference might be caused by a small facilitation in the Overlap condition because it shares its CV structure with the Baseline condition. If this hypothesis is true, the difference between the Anagram and Baseline conditions should be maximal in the earlier stages of production, when the CV frame is still being built/retrieved. Specifically, the effect should show the largest difference in RTs and initial inter keystroke intervals (IKIs) and disappear on later IKIs in the Anagram condition. Experiment 2 was designed to answer these questions. We used the interference effect in durations in the Overlap condition in Exp 1 (which was less consistent than the Anagram condition), to estimate the sample size needed to replicate the effect at $\alpha = 0.05$ with a power of .80. The estimated sample size was 20 (note that the modality differences might have an effect, but this was the closest effect size available to make a reasonable estimation). We thus targeted a sample size of 24 (allowing for potential attrition) to conduct the same experiment, this time instructing participants to type their responses.

Experiment 2

Methods

Participants Twenty-four English-speaking undergraduates (11 female, 13 male; mean age = 19.71 (SD = 1.23)) participated for course credit. None had participated in Experiment 1. Data from one participant was excluded from analysis due to failure to follow task instructions.

Stimuli Stimuli were the same as those used in Experiment 1.

Procedure Procedures were similar to those used in Experiment 1. The experiment was displayed on a 20.41×12.75-inch Dell monitor approximately 24 inches in front of participants. RTs, durations, and IKIs were recorded on a millisecond-accurate Empirisoft DirectIN PCB v2016 keyboard. Participants were instructed to type the name of each image as quickly and as accurately as possible. To submit their response, participants simultaneously pressed the left and right shift keys to reset their hand positions in preparation for the next trial. In each trial, a fixation cross was presented in the center of the screen for 700 ms, followed by the presentation of the image. Participants then had 1500 ms to finish typing their response. If participants did not complete their response within 1500 ms, they were prompted on screen to press both shift keys to proceed to the next trial. This experiment took approximately 40 minutes to complete, including two short breaks.

Analysis Trials were excluded according to the same criteria as Experiment 1, except the minimum RT was reduced to 100 ms to reflect the shorter RTs in typing. Analyses of RTs and durations were identical to Experiment 1. Three IKIs were also analyzed: between the first and second (IKI1), second and third (IKI2), and third and fourth (IKI3) keypresses. Each IKI was analyzed separately in the same way as the durations. Unlike RTs and durations, we did not have clear a priori predictions regarding effects on specific IKIs. All IKI analyses were thus corrected for multiple comparisons.

Results & Discussion

Because we were interested in analyzing the timing of individual keystrokes, the exclusion of all trials that contained an incorrect letter was unavoidable, even if participants had efficiently corrected that letter. Together with uncorrected errors, these trials comprised 9.9%, 9.1%, and 10.7% of trials in Baseline, Overlap, and Anagram conditions, respectively, and did not differ between conditions. With an additional 1.5% of trials excluded based on the criteria defined under Analysis, 5,547 target trials were included in the analyses.

Figure 3 shows the RTs, durations, and IKIs for the target words in the three conditions in Exp 2. As in Exp 1, durations were significantly longer for both Overlap and Anagram conditions compared to the baseline (*M* = 6.57 ms, *p* = .042, and *M* = 13.54 ms, *p* < .001, respectively). This time, the direct comparison between the two conditions also revealed significantly longer durations in Anagram condition (*M* = −6.98 ms, *p* = .030). As expected, cleaner RT effects emerged in Exp 2. Both Overlap and Anagram conditions showed significantly longer RTs compared to Baseline (*M* = 14.03 ms, *p* = .009, and *M* = 23.71 ms, *p* < .001, respectively). The difference between the Overlap and Anagram conditions was marginal (*M* = −9.68 ms, *p* = .071). The results of the LMEM analyses replicated the effect of Overlap and Anagram against Baseline, but the comparisons between Overlap and Anagram were non-significant on RTs (*p* = .218) and marginal on durations (*p* = .084). After correcting for three comparisons, IKI1 was only significantly longer in Anagram compared to Baseline conditions (*M* = 3.98 ms, corrected *p* = .033). IKI2, on the other hand, was only significantly longer for Overlap vs. Baseline (*M* = 4.72 ms, corrected *p* = .012). IKI3 was not significantly different in any comparisons (all p-values > .1). LMEM results were identical.

To summarize, the results of Exp 2 largely replicated those of Exp 1 and cleaned up the discrepancy in the RT effects found in Exp 1. A robust interference effect was found on both RTs and durations on targets when they were produced in the presence of words that shared segments with them, irrespective of whether the shared segments were in the same or different positions, i.e., in both Overlap and Anagram conditions compared to Baseline. Interestingly, Experiment 2 also replicated the unreliability of the difference between Overlap and Anagram conditions, which was significant in one method of analysis but not the other. It must be noted though, that the pattern of results was identical in both experiments and consistent across RT and duration measures; there was always a slightly *larger* interference effect on Anagram compared to Overlap conditions, but the effect in individual measures and experiments was not robust. Note again that because of the direction of the effect, the interpretation of the main finding remains unambiguous: segmental interference is not limited to shared segments in the same position. Nevertheless, the IKI analyses may shed some light on the slightly larger interference effect in Anagram vs. Overlap conditions: interference in the Anagram condition was evident on IKI1, whereas interference in the Overlap condition was only evident on IKI2. This is consistent with greater difficulty in filling earlier rather than later slots in the Anagram condition, which, in turn, may reflect difficulty in building the CV frame in this condition. In the Overlap condition, the CV structure is, by definition, shared between target and competitor, and could be built/retrieved much faster on each trial.

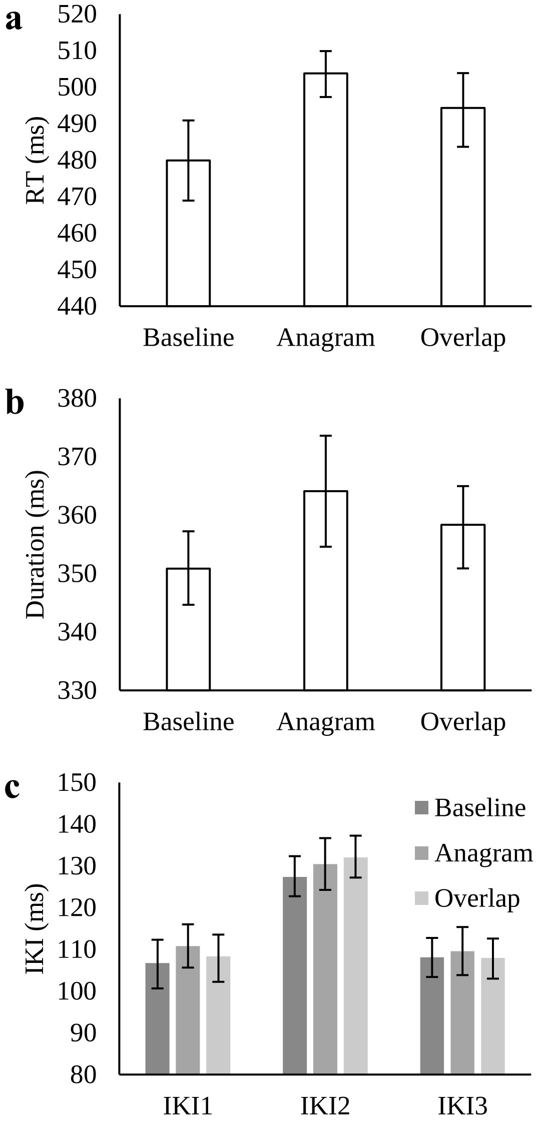

Figure 3: Mean (a) RTs, (b) durations, and (c) IKIs for responses to target items in Baseline, Anagram, and Overlap conditions across participants in Experiment 2. Error bars represent 95% CIs not corrected for between-subject variability.

General Discussion

In two experiments, we employed the segmental interference effect to test whether segments cross-activate similar segments in different syllabic positions or not. The results suggest that they do. Both "flow" and "wolf" interfered with the production of "glow". This finding supports a layer in the production system in which segments are represented independent of their context. Moreover, it suggests that such segmental representations can in fact activate words that contain similar segments, irrespective of the position of those segments. Such segments gain positional representation through links with specific positions in individual words (Fig. 1b; e.g., Houghton, 2018). This finding has important implications for models of spelling, and reading, most of which only represent segments in a position-dependent manner. Even if this was simply a pragmatic choice, those models would not be able to simulate the current findings. More generally, the current data suggest that capturing the true scope of interference effects requires models that posit position-independent segmental representations, which can feed back to lexical items before they are bound to specific positions in the syllabic frame.

There was no evidence that the interference effect induced by Anagram competitors was smaller than that induced by Overlap competitors. If anything, the direction of the effect was consistently towards a *larger* interference effect in the Anagram condition (a consistent pattern across both RT and duration measures and in both experiments, but not statistically significant in all tests). It is possible that this difference is due to the fact that the competitors in the Anagram condition, unlike the Overlap condition, do not share the CV structure with the target, so the switch between CV frames adds some cost. This hypothesis is supported by an earlier locus of interference in IKI (IKI1 in Anagram and IKI2 in Overlap). Note, however, that the difference in the CV structure cannot be the main source of interference in the Anagram condition, because the competitors in the control condition also do not share their CV structure with the target. If that were the main source of interference, there should have been no significant differences between Anagram and Baseline conditions.

The convergence of results in handwriting and typing points to some level of modality-independence for the effect. However, extrapolation to spoken production requires testing the effect in the spoken modality. The findings so far are inconsistent: facilitation in picture naming using phonologically related auditory primes has been shown to be position-specific (Gagnon & Sawusch, 1989), but as discussed in the Introduction, facilitatory effects of phonological overlap tend to have a strategic component. In line with this reasoning, Gagnon and Sawusch (1989) showed that priming-related facilitation was found to depend on speaker-related variables. Moreover, they did not replicate the position-specificity of segments when a different production task was used. The paradigm tested in the current work offers a solid alternative for investigating whether the position-independence of segmental representations extends to spoken production.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Breining, B. L., Nozari, N., & Rapp, B. (2019). Learning in complex, multi-component cognitive systems: Different learning challenges within the same system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(6), 1093–1106.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, *43*(2), 182–216.

Gagnon, D. A., & Sawusch, J. R. (1989). Converging evidence on the nature of the segmental representation underlying spoken word recognition. *The Journal of the Acoustical Society of America*, *86*(S1), S99–S100.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 9, pp. 133–177). Academic Press.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*(3), 491–528.

Houghton, G. (2018). Action and perception in literacy: A common-code for spelling and reading. *Psychological Review*, *125*(1), 83–116.

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology*, *20*(2), 115–162.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–146). Wiley.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–38.

Nozari, N., Freund, M., Breining, B., Rapp, B., & Gordon, B. (2016). Cognitive control during selection and repair in word production. *Language, Cognition and Neuroscience*, *31*(7), 886–903.

O'Séaghdha, P. G., & Frazer, A. K. (2014). The exception does not rule: Attention constrains form preparation in word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 797–810.

R Core Team. (2019). *R: A language and environment for statistical computing* (Version 3.6.1) [Computer software]. R Foundation for Statistical Computing.

Wickelgran, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, *76*(1), 1–15.