EFFECTS OF PHONEME REPETITION IN SPOKEN UTTERANCE GENERATION

Markus F. Damian and Nicolas Dumay

Department of Experimental Psychology, University of Bristol, UK
m.damian@bristol.ac.uk, n.dumay@bristol.ac.uk

ABSTRACT

The degree of phonological advance planning in spoken production was investigated with a paradigm in which speakers performed speeded naming responses to coloured line drawings of objects. Colours and object names were chosen such that a phoneme matched, or mismatched, between adjective and noun. A facilitatory effect of repeated phoneme was demonstrated, which was found not only when the phoneme occupied the word-initial position ("green goat"), but also in the central ("black pan") or word-final ("black monk") position. These results imply that speakers planned the phonological content of the entire phrase before starting their articulation. A facilitatory effect was additionally found when the repeated phoneme occupied a different position within each word ("green flag"). The latter result suggests that the spoken production system represents segments independently of their position within a word.

Keywords: spoken production, advance planning, phonological encoding

1. INTRODUCTION

A central issue in research on spoken production pertains to the extent to which speakers plan ahead. Various studies suggest that at higher planning levels (e.g., conceptual or syntactic) the extent of planning is larger than at lower levels (i.e., phonological). For instance, phoneme exchanges in speech errors typically occur between adjacent words, whereas semantic exchanges are largely unconstrained by linear position in an utterance [4]. In experimental settings designed to investigating the scope of advance planning, the issue can be reformulated as: what portion of an utterance has been planned at the time pronunciation is initiated.

The work reported below specifically looks at planning at the phonological level. The extent of planning at this level remains controversial. "Minimalist" accounts of phonological planning stipulate that only a negligible degree of "buffering" is taking place prior to the start of articulation, quite possibly less than a single word [8]. "Maximalist" accounts hypothesize that phonological planning is substantial and may be largely driven by syntactic constraints [3]. Perhaps the most popular account, the "phonological word hypothesis" [5, 12, 13], postulates an intermediate degree of planning. According to this account, advance planning comprises at minimum a single phonological word, defined minimally as a stressed foot, and maximally as a lexical word together with associated unstressed function words such as determiners, auxiliaries, prepositions and conjunctions.

One way of evaluating these scenarios is to ask speakers to produce multi-word utterances, such as adjective-noun phrases, and to manipulate the phonological overlap between the two constituents. In the experiments described below, we asked speakers to name coloured line drawings of common objects under time pressure, with the colour adjective and picture name selected such that a single phoneme was either repeated ("green goat", "red rake") or not ("red goat", "green rake"). The logic is that if phoneme repetition affects latencies to initiate the utterance, speakers must have planned both constituents before starting their response.

Effects of phoneme repetition on the psycholinguistic aspects of speech production have rarely been investigated, and the existing results do not allow clear predictions for the situation used in the present study. For instance, in the speeded recitation of sequences of four CVC words, the repetition of the same final consonant from item to item ("pick tuck puck tick") accelerates the articulation rate, whereas the repetition of the same initial consonant produces interference ("pick pun puck pin") [10]. In any case, whether it is facilitatory or inhibitory, an effect of phoneme repetition on the response latencies in the experiments below would indicate planning of the entire phrase.

On the phonetic front, the existing studies have essentially looked at the specific case of "gemination", where the repeated phonemes are adjacent ones ("green nest"). However, because gemination

is characterized by articulatory features constrained primarily by the close proximity of the two phonemes, we do know whether phoneme repetition on its own right has an impact on the realization of segments. To address this question, the first experiment included also measures of response duration. Here, we predicted that under the assumption that planning was completed prior to response initiation, durations should be unaffected by the relatedness manipulation.

2. EXPERIMENT 1

12 speakers were asked to name coloured pictures presented on a computer monitor with an adjective-noun phrase, and responses were digitised on-line. 20 line drawings of objects with monosyllabic names served as targets, and were presented in one of four colours (red, blue, green, pink). Each target was paired with a color that either matched ("green gun", related) or mismatched ("red gun", unrelated) regarding the initial segment. Each target and each colour appeared in both the "match" and the "mismatch" condition. In addition to these 40 critical trials, we included 80 filler trials in which the target pictures were paired with phonologically unrelated colour names, in order to reduce the likelihood that participants developed expectancies about related pairings.

Adjective onset, noun onset and response offset were determined for each individual trial in a waveform editor. From these, average latencies and durations for each constituent were computed. The results are shown in Table 1.

Analyses of variance (ANOVAs) by participants ($F1$) and items ($F2$) performed on onset latencies showed a significant facilitatory effect of repeated phoneme, $F1(1, 11) = 19.58$, $p = .001$; $F2(1, 19) = 5.20$, $p = .034$, and similarly significant effects of relatedness on noun onset, and utterance offset. By contrast, no significant effects of relatedness were found on overall duration, nor on duration of adjective and noun separately. Error percentages in the two conditions did not differ significantly. These results suggest that prior to initiation of articulation, speakers phonologically planned both adjective and noun. Note that the fact that response execution was equivalent for the related and unrelated condition further strengthens the conclusion that planning took place before articulation began.

Table 1. Experiment 1 – Onset of adjective, noun and offset, and corresponding durations, dependent on relatedness. Standard deviations in brackets.

	Related	Unrelated	Priming
$Onset_{adj}$	823 (38)	870 (45)	+47
$Onset_{noun}$	1130 (48)	1170 (45)	+40
Offset	1599 (71)	1645 (78)	+46
$Duration_{adj}$	301 (16)	300 (16)	-1
$Duration_{noun}$	475 (27)	477 (25)	+2
$Duration_{overall}$	775 (40)	775 (38)	0

3. EXPERIMENT 2

Previous studies investigating speech errors have demonstrated an "initialness effect", i.e., a prominence of word-initial phonemes involved in errors [6]. For instance, consonant slips exhibit a strong tendency to occur in syllable onsets rather than in codas [11]. One way of accounting for this phenomenon is to postulate that initial elements exhibit particular salience within an utterance. For example, it has been claimed that "initial sounds slip a lot because they are, in general, easy to retrieve – or, to use activation terms, they become highly activated quickly" [1]. Regarding the findings from coloured object naming outlined in our first experiment, this raises the issue of whether the effect of repeated phonemes is specific to the initial position of adjectives and nouns, or whether it can also be obtained in other positions.

We addressed this issue in a second experiment. 24 participants were again asked to name line drawings of objects with monosyllabic names, presented in one of three colours (red, green, black). For the "initial overlap" condition, twelve objects were selected as targets, and paired with colours such that the initial phoneme of adjective overlapped ("black boat") or not ("green boat"). For the "central overlap" condition, twelve further objects were selected such that the central vowel of color adjective and name matched ("black pan"), or not ("red pan"). Finally, for a "final overlap" condition, twelve different objects were selected such that the final consonant of color adjective and name matched ("black monk") or mismatched ("red monk").

The results are shown in Table 2. A facilitatory effect of phoneme overlap, which was largely independent of position of the repeated phoneme within the utterance, was obtained. An analysis of

variance, with relatedness and overlap position as the variables, showed a significant effect of Relatedness, $F1(1, 21) = 8.39$, $p = .009$, $F2(1, 33) = 14.00$, $p < .001$, with latencies being 22 ms faster in the related than in the unrelated condition. The effect of overlap position was significant in the analysis by participants, $F1(2, 42) = 4.71$, $MSE = 11830$, $p = .014$, but not by items, $F2 = 1.52$, $p = .234$. Importantly, the interaction between Relatedness and Overlap position was not significant, $F1$ and $F2 < 1$. Error percentages in the two conditions did not differ significantly. In sum, the facilitatory effect of phoneme repetition between colour adjective and picture name is largely independent of its position within each word.

4. EXPERIMENT 3

The experiments reported above demonstrate that phonological planning of the adjective-noun utterance results in a facilitatory effect when a phoneme is repeated, and that this effect is independent of the position of overlap within each word. This raises questions about how within the spoken production system, position of a segment is represented. Most computational attempts to model phonological encoding in speaking have avoided the issue by choosing to model words that have no repeated elements. The prominent model of speech production brought forward by Dell and colleagues [1, 2] implements a so-called "slot coding" scheme in which phonemes are coded separately for within-word position. This implies that, e.g., the segment "g" within a bank of onset consonants is entirely unrelated from the same segment "g" when occurring in the bank of coda consonants. This coding scheme allows for an interesting prediction: the effect of repeated phoneme demonstrated above should be position-specific, i.e., it should only be observed if the critical segments occur within the same position in both words. Hence, relative to an unrelated condition, priming is predicted for, e.g., "green goat", but not for "green flag", in which the overlapping segment occupies the initial position in the adjective, but the final position in the noun.

We assessed this issue in a third experiment. 24 participants were again asked to name line drawings of objects with monosyllabic names, presented in one of three colours (blue, green, pink). For the "initial overlap" condition, fifteen line drawings were selected as targets, and were paired with the colours such that the initial phoneme either matched ("pink pen") or mismatched ("blue pen"). For the "initial/final overlap" condition, fifteen further objects were selected and paired with colours such that the first phoneme of the adjective, and the final phoneme of the name matched ("blue crab") or mismatched ("green crab").

The results are shown in Table 3. As in the first two experiments, word-initial overlap resulted in a sizeable facilitatory effect. By comparison, the effect in the initial/final condition was numerically reduced, but still clearly present. An Anova conducted on the data, with relatedness and overlap position as the variables, showed a significant effect of relatedness, $F1(1, 23) = 14.93$, $p < .001$, $F2(1, 28) = 6.91$, $p = .014$. The effect of overlap position was significant in the analysis by participants, $F1(1, 23) = 12.33$, $p = .002$, but not by items, $F2(1, 28) = 2.36$, $p = .136$. Importantly, no interaction between the two factors was found, $F1(1, 23) = 1.10$, $p = .305$, $F2 < 1$. This implies that statistically, the repeated phoneme effect does not depend on whether the critical element occurs in the same, or a different, position within the words. Error percentages in the two conditions did not differ significantly.

5. DISCUSSION

The results can be summarised as follows. In coloured object naming, latencies are faster when phonemes in the adjective overlap with those in the noun. Experiment 1, with word-initial overlap, showed that this effect manifests itself exclusively

Table 2. Experiment 2 – Response onset, dependent on overlap position and relatedness. Standard deviations in brackets.

Overlap Position	Related	Unrelated	Priming
Initial	771 (133)	796 (122)	25
Central	740 (105)	762 (99)	22
Final	755 (120)	775 (112)	20

Table 3. Experiment 3 - Response onset, dependent on overlap position, and relatedness. Standard deviations in brackets.

Overlap Position	Related	Unrelated	Priming
Initial	768 (99)	801 (105)	33
Initial/ Final	798 (98)	815 (101)	17

in response latencies, but not in response durations. This implies that speakers apparently planned the entire phrase before initiating their response. The result is seemingly incompatible with "minimalist" accounts of speech planning, which stipulate that only minimal buffering occurs before speakers start a response. However, the results are also problematic for the "phonological word hypothesis", which assumes that a single phonological word constitutes the minimum possible degree of planning. Speakers in our experiments produced utterances consisting of two phonological words (unless it is assumed that because all utterances were disyllabic, they were treated as a single phonological word), and the effects of phoneme repetition suggest that the entire phrase was planned. From the perspective of theories of phonological advance planning in speaking, the results therefore confirm recent claims that the degree of planning at that level is more extensive than previously believed [9].

Experiment 2 additionally demonstrated that the effect is not contingent on the repeated phoneme occupying the initial position within adjective and noun. This finding is nontrivial because studies on speech errors have pointed out a particular vulnerability of the word-initial position for segment exchanges. Our results, by contrast, suggest that repeated phonemes accrue a facilitatory effect on latencies even when occurring at non-initial positions within the utterance. Indeed, Experiment 3 showed that the repeated phoneme does not necessarily need to occupy the same position within the two words - even a condition in which a phoneme occurring at the beginning of the adjective is repeated at the end of the noun, induces facilitation. The latter finding has important implications for computational models of spoken production, and specifically for the way in which phonemes are assigned to word positions. Our results are difficult to accommodate in a slot coding scheme, as implemented in the model advocated by Dell and colleagues [1-2], because one and the same phoneme which occurs in different word positions ("banks") is treated as two entirely separate instances, and hence it is difficult to see how the obtained facilitation could emerge.

On the other hand, the computational framework WEAVER [5, 7] will likely find it more straightforward to account for our results because this model does not rely on a specific coding for word position. Instead, a target word form activates in parallel all associated phonemes by means of links which specify the linear position within the word. As opposed to "slot coding" schemas, in this model each segment is represented only once, but it can in principle be accessed more than once within an utterance by means of several position-specific links pointing to it. To our knowledge, the situation of a particular phoneme appearing repeatedly within a word (or utterance) has not been explicitly modeled. However, it seems plausible that the repeated segment will accrue additional activation compared to the unrelated case, resulting in faster phonological encoding. Importantly, this will occur independently of position of repetition within the utterance. In this way, our findings provide important constraints on the structure of computational models of phonological encoding.

6. REFERENCES

[1] Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.

[2] Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*, 801-838.

[3] Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review, 100*, 233-253.

[4] Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, pp. 133-177). New York: Academic Press.

[5] Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

[6] MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, 8, 323-350.

[7] Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*, 249-284.

[8] Santiago, J., MacKay, D. G., Palma, A., & Rho, C. (2000). Sequential activation processes in producing words and syllables: evidence from picture naming. *Language and Cognitive Processes, 15*, 1-44.

[9] Schnur, T. T., Costa, A., & Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of Psycholinguistic Research, 35*(2), 189-213.

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

[11] Stemberger, J. P. (1982). *The lexicon in a model of language production.* Unpublished doctoral dissertation, University of California, San Diego.

[12] Wheeldon, L. (2000). Generating prosodic structure. In L. R. Wheeldon (Ed.), *Aspects of language production* (pp. 249-274): Hove: Psychology Press.

[13] Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37*, 356-381.