eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Serial order in word form retrieval: New insights from the auditory picture-word

interference task

Carolyn Wilshire and Sunita Singh

School of Psychology, Victoria University of Wellington, New Zealand

Catherine Tattersall

Department of Human Communication Sciences, University of Sheffield, United

Kingdom

Correspondence to: Carolyn Wilshire, School of Psychology, Victoria University of

Wellington, P.O. Box 600, Wellington, New Zealand.

E-mail: Carolyn.Wilshire@vuw.ac.nz

Phone: +64 4 463 6036

Fax: +64 4 463 5402

Abstract

One important theoretical question about word production concerns whether the phonemes of a word are retrieved in parallel or in sequential order. To address this question, Meyer and Schriefers (1991) used an auditory picture-word interference task, and manipulated the position of the phonemes shared between distractor and target picture. They found that begin-related distractors (e.g., *boat-bone*) facilitated naming times when presented within 150ms before or after the picture, whereas end-related distractors (e.g., *cone-bone*) were effective only if presented within 150ms after the picture. This suggested that the word's end phonemes were activated later than beginning ones. However, it is unclear whether these effects genuinely reflect facilitation at the level of phonological retrieval. In this study, we examine later distractor presentation onsets, where distractors have little opportunity to influence earlier, lexical selection processes. At the latest onset tested, end-related – but not begin-related - distractors significantly facilitated naming. We conclude that late-presented distractors do indeed influence phonological encoding, and that their asymmetric effects support a sequential model of phoneme retrieval.

Serial order in word form retrieval: The auditory picture-word interference task revisited

Current theories propose that word retrieval involves at least two major stages: *lexical selection*, where the speaker selects a word that matches the desired concept, and *phonological retrieval*, where s/he then retrieves its phonological form. This information provides the input to subsequent, articulatory-motor programming processes (see for example, Dell, 1986, 1988; Foygel & Dell, 2000; Levelt, Roelofs & Meyer, 1999; Rapp & Goldrick, 2000; Roelofs, 1997, 2004). Most contemporary models propose that the phonemes of a word are retrieved simultaneously, irrespective of their position. It is only at subsequent, articulatory-motor stages that operations become sequential (Levelt et al., 1999; Roelofs, 1997; Roelofs, 2004; Shattuck-Hufnagel, 1992). For example, Roelofs and colleagues propose that all phonemes of a morpheme are activated in parallel, accompanied by number labels to identify their position (this is called *segmental spellout*). The phoneme string is then converted into a series of articulatory motor commands by a process that operates sequentially (called *syllabification and prosodification*; Levelt et al., 1999; Roelofs, 1997, 2004). However, a contrasting possibility is that phoneme retrieval is itself sequential: that is, the word's early phonemes initially receive more activation than later ones, and as retrieval unfolds, later phonemes receive progressively more activation (Dell, Juliano & Govindjee, 1993; Hartley & Houghton, 1996; Houghton, 1990; Sevald & Dell, 1994; Vousden, Brown, & Harley, 2000; see also Dell, 1986, for an intermediate view).

One paradigm that might help adjudicate between these two possibilities is the *auditory picture-word interference task*, where pictures to be named are accompanied by auditory distractor words, which the participant simply ignores. Pictures are named faster if the distractor shares some of the target's phonemes (e.g., *boat-bone*), than when it does not (e.g., *cat-bone;* Abel et al., 2009; Damian, Bowers, Stadthagen-Gonzalez & Spalek, 2010;

Damian & Martin, 1999; de Zubicaray & McMahon, 2009; Jescheniak & Schriefers, 2001; Schriefers, Meyer & Levelt, 1990; Starreveld, 2000). This *phonological facilitation effect* has been observed when the distractor is presented anywhere from 150ms before the picture to 200ms after it, and sometimes at even earlier presentations (300ms before: Jescheniak & Schriefers, 2001; Starreveld, 2000).

Most studies reporting this effect have used distractor pairs that shared the same onset phonemes (e.g., *boat-bone*). Few have varied the position of the shared phonemes in a way that might allow us to explore serial order effects. One exception is the seminal paper of Meyer and Schriefers (1991, henceforth MS1991), which compared the effects of *begin-related* distractors (those sharing the first two phonemes, or for bisyllabic words, the first syllable) and *end-related* distractors (those sharing the final two phonemes, or for bisyllabic words, the second syllable). In their Experiment 3, distractor words were presented so that the critical overlapping phonemes occurred at one of four onsets: 300ms before the picture; 150ms before the picture; simultaneously with the picture; or 150 ms after the picture. Begin-related distractors facilitated naming relative to unrelated ones at all but the earliest of these onsets. End-related distractors did so only at the two latest onsets. They concluded that the distractor "preactivates" those phonemes it shares with the target, thereby reducing the time required to retrieve them for production. Further, since begin-related distractors are effective at earlier presentation times than end-related ones, the onset phonemes in the target word must become available before later ones (see also Damian et al., 2010; Meyer & van der Meulen, 2000; Schriefers, 1999).

However, this conclusion is undermined if phonological distractors also facilitate *lexical selection* (Damian et al., 2010; Damian & Martin, 1999; Jescheniak & Schriefers, 1998, 2001; Levelt et al., 1999; Roelofs, Meyer, & Levelt, 1996). When an auditory word is presented, it likely activates the lexical representations of similar sounding words,

particularly those with similar onset phonemes (Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994). A phonologically related distractor might therefore "preactivate" the target lexical representation, allowing it to be selected more rapidly. The fact that begin-related phonological facilitation can be observed even when distractors are presented a full 300 ms before the picture – and well before phonological retrieval would be expected to begin – supports this possibility. Further, in an fMRI study, De Zubicaray and McMahon (2009) found that begin-related distractors presented 200ms before the picture yielded reduced signal activation not only in the left posterior superior temporal gyrus - a region associated with phonological processing - but also in the middle temporal gyrus, a region argued to play a key role in lexical selection (Acheson, Hamidi, Binder & Postle, 2011; Schwartz et al., 2009; see also Indefrey & Levelt, 2004; Indefrey, 2011). In the context of MS1991, then, begin-related distractors may simply be more effective at facilitating lexical selection than end-related distractors, and this type of facilitation may be more robust to variations in distractor onset time. Indeed, in MS1991, not only did the facilitatory effect of begin-related distractors occur at earlier onsets than that for end-related ones, but also its magnitude was considerably greater. Further consistent with the lexical view, there was no presentation onset at which end-related distractors were more effective than begin-related ones.

These concerns over a study over two decades old may seem obscure, but they have significantly impacted upon the theoretical landscape in this field. Suspicion of evidence from picture-word studies has contributed to the prevailing view that processing does not become sequential until the articulatory-motor planning stage (e.g., Damian & Dumay, 2009; Roelofs, 2004). Despite this view, however, there has actually been little direct examination of the precise locus of different types of phonological facilitation effects within the picture-word paradigm. In this study, we attempt to remedy this situation. We explore the effect of

two different types of distractors – begin-related and end-related – across four well-spaced presentation onsets: -200ms (before the picture), 0ms, 200ms or 400ms. We reasoned that, the later a distractor word is presented, the less likely it will be to impact upon lexical selection. Current estimates of lexical selection time for pictures presented isolation indicate it occurs within the first 200ms (see, e.g., Strijkers & Costa, 2011), and even in the context of an interference paradigm, it is likely to be complete well within 300ms. If a distractor word is presented so late that the critical overlapping phonemes are not processed until well after this time, its effects are more likely to be attributable to processes occurring further downstream. Figure 1 illustrates this scenario graphically. In particular, if we find end-related facilitation in the absence of begin-related facilitation at such late onsets, this would support a sequential model of phoneme retrieval. If, however, both - or neither – type of facilitation occurs, then this would favour a parallel model of phonological retrieval.

Insert Figure 1 here

To maximise our chances of observing serial order effects, all targets were bisyllabic. Further, to reduce the predictability of the targets, we used a large pool of 96 target pictures (MS1991 used 25). Also, to minimise the need for extensive training on picture names, the pictures were pre-piloted to ensure they had high name agreement. Finally, unlike in MS1991, where different participant groups completed the different presentation onsets, all participants completed all onsets, in a crossed within-subjects design; we alternated the identity of the auditory distractors paired with each picture, to minimise the chances of distractors acting as episodic cues as to the target.

Method

Participants

Twenty four undergraduate Psychology students (17 females), aged from 18 to 24 years (*M*=20.08 years) participated for course credit. All had normal vision and hearing and were native English speakers.

Materials

The 96 target words were all bisyllabic nouns with Celex lemma frequencies ranging from 0 to 409 tokens per million (geometric mean 9.88), with four to seven phonemes (mean 4.72). Their pictures were coloured line drawings and photographs, all of which yielded name agreement of 80% or more when piloted on a group of 70 speakers of New Zealand English of varying ages. For each target word, two exemplars of each of the following distractor word types were selected: a) begin-related distractors, in which the distractor word generally shared the target word's first syllable and stress pattern (e.g., *turkey - turnip*)[1]; b) begin-unrelated distractors, which were obtained by randomly reassigning the begin-related distractors to different targets; c) end-related distractors, in which the distractor word generally shared the target's second syllable and stress pattern (e.g., *turkey - hockey*)[2]; and d) end-unrelated distractors, obtained by randomly reassigning the end-related distractors to different targets. There were no significant overall frequency differences between the four types of distractors, nor between the two exemplars of each distractor type. All distractors were digitally recorded by a native speaker of English.

Design

Each picture was presented in each of the distractor conditions at four different stimulus onset asynchronies (SOAs): -200 ms, (distractor before picture); 0 ms, +200 ms (distractor after picture); and +400 ms. The two exemplars of each distractor type were alternated across SOAs (e.g. for *lemon*, the begin-related distractor used at –200 ms and +400 ms was *leather* and the one used at other SOAs was *leopard*). [3] SOA timings were always calculated with respect to the point of overlap between critical syllable of the distractor and

picture. So, for begin-related distractors, the SOA was the interval between picture onset and that of the distractor's first syllable; for end-related distractors, it was the interval from picture onset to the onset of the distractor's *second* syllable (calculated individually for each distractor word). This aspect of the design is illustrated in Figure 1. For the control, unrelated conditions, distractors were presented at exactly the same SOA as for the corresponding related condition.

Since each target picture appeared in 16 conditions (four different distractor conditions x four SOAs) and there were 96 target pictures, there were 1536 trials in total. These trials were organised into 16 blocks of 96 trials each, with each target picture appearing once in each block, and each block containing a roughly equal number of examples of the various condition by SOA combinations. Within each of the experimental blocks, the order of trials was pseudorandomised, with the limitation that no more than two successive trials could feature the same distractor condition, or could include targets with the same onset. Each of the 24 participants completed half of all the total possible trials (eight blocks, 768 trials), yielding a total of 18,432 trials across the entire experiment.

Procedure

Stimuli were presented to participants on a Macintosh computer using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were randomly assigned to one of six different groups; each received a different combination of blocks in each of the sessions. Testing was spread over two testing sessions spaced at least three days apart. Prior to the first session, all participants completed a computerised naming task featuring all 96 stimulus pictures, and if they gave a non-target word, the target was provided. In the experimental sessions, participants were not corrected.

Participants were instructed to name the pictures and ignore the distractor words. The participant pressed a key to start the trial, and 700ms later, the picture appeared. The onset of

the distractor varied according to SOA condition. The picture remained on the screen throughout the naming attempt. All sessions were recorded digitally. Naming latencies for responses that were correct on the first attempt were manually measured using digital sound analysis software, from the onset of the target picture to that of the naming response.

Results

A small proportion of trials (0.38%) were lost due to recording or software failures. Participants responded correctly to 96.9% of begin-related trials, and 96.1% of begin-unrelated trials. For the end-related condition, these percentages were 96.1% and 96.7% respectively.

Prior to analysis of latencies, we removed all trials incorrect on the first attempt (4.1% of trials) and all trials involving a target that was failed more than twice by that participant (1.0% of trials). Then outliers, defined as latencies more than 2.5 standard deviations above the participant's mean, were removed (2.9% of trials). Latency data were then submitted to Linear Mixed Effects analysis using the SAS Proc Mixed procedure. The model was estimated using Restricted maximum likelihood. The fixed effects were: relatedness (related, unrelated), position of overlapping phonemes (begin vs. end), SOA, and all possible interactions. Both participant and target name were entered as crossed random effects, and for the participant random effect, slopes were specified with respect to each of the three main effects (but not for the item random effect, since in no instance did the inclusion of slopes significantly improve model fit). The covariance structure, specified prior to analysis, was compound symmetry for both random effects. Degrees of freedom were calculated using the Satterthwaite approximation. F values reported in the text are the Type 3 hypothesis tests for each of the specified fixed effects and for the contrasts of interest (performed using the contrast statement in Proc Mixed). To allow for comparison of effect sizes across studies, values for generalized eta squared are also provided, calculated in the conventional manner

from data collapsed across items (referred to in the text as $\eta_G^2(F_1)$). Finally, for the key contrasts, Bayes Factors were estimated using the approximation suggested by Johnson (2005).

Table 1 shows the mean naming latencies to correct responses across conditions. Figure 1 shows the mean differences in latency between each related condition and its respective control. There was a significant main effect of relatedness, $F(1,84.2)= 29.86$, $p<0.0001$, $\eta_G^2(F_1)=0.13$, and a significant interaction between relatedness and position of shared phonemes, $F(1,17000)= 12.90$, $p<0.001$, $\eta_G^2(F_1)= 0.03$. Finally, this relatedness by position interaction was significantly modulated by SOA, $F(3,17000)=7.90$, $p<0.0001$, $\eta_G^2(F_1)= 0.03$. The effect of begin relatedness was significant at the -200ms SOA ($p<.05$, $BF_{01}=5.19$), at 0ms ($p<.0001$, $BF_{01} > 10^{12}$) and at +200ms ($p<.0001$, $BF_{01}> 10^6$), but not at +400ms ($p=.841$, ns, $BF_{01}=2.09$). The effect of end-relatedness was significant at 0ms ($p<.01$, $BF_{01}=4.67$), and 400ms ($p<0.01$, $BF_{01}=9.69$), but not at -200ms ($p=.055$, ns, $BF_{01}=0.21$) or +200ms ($p=.198$, ns, $BF_{01}=0.08$). Finally, in order to directly compare the effects of begin- versus end-relatedness at each SOA, we also performed contrasts to test the interaction between relatedness and position of shared phonemes at each individual SOA. There was a significant relatedness by position interaction at all but the earliest SOA (-200ms: $p=.12$, ns, $BF_{01}=0.29$; 0ms: $p < .0001$, $BF_{01}=409.25$; +200ms: $p > .001$, $BF_{01}=190.88$; +400ms: $p = .041$, $BF_{01}=1.39$).

Insert Table 1 and Figure 2 here

Discussion

In this study, we observed large and statistically reliable effects of begin-relatedness at SOAs of -200ms, 0ms and +200ms, but not at our later presentation onset of 400ms (see also Damian, 2003, for a similar finding). There was also a statistically reliable but smaller effect of end-relatedness at 0ms, and at our late, 400ms onset (but not at the +200ms onset).

Of greatest interest is the effect at the late, 400ms SOA, where end-related distractors were reliably more facilitatory than begin-related ones. According to even the most generous estimates, lexical selection here is likely to be complete well before the critical overlapping phonemes in the distractor are processed. These effects are therefore far more likely to be attributable to processes occurring further downstream. Also, the *greater* facilitation for end-related than for begin-related distractors at this SOA is difficult to reconcile with a lexical-level account. While it is possible that auditory words may activate other words that share the same end phonemes (as recently argued by Damian et al., 2010), to our knowledge there are no theories that would allow for such effects in the absence of comparable or greater begin-sharing effects. Therefore, we conclude that, at least under certain conditions, phonologically related distractors can indeed facilitate phonological retrieval, but such effects may only be clearly evident at very late presentation onsets.

Importantly, the fact that the facilitation observed here was from purely auditory stimuli – that were not themselves ever produced by the speaker – makes it unlikely that they arose at subsequent, motor-articulatory processing stages. Roelofs and colleagues allow for the possibility that articulatory-motor respresentations of syllables may be pre-prepared or even primed from previous use, but there is no mechanism that enables then to become activated directly from purely auditory input (Roelofs, 1997, 2004).

From a theoretical point of view, the conclusion we draw that there must be a time at which when a word's beginning phonemes are fully retrieved but its end phonemes (in this case, its second syllable) are not yet so. This supports the view that phonological retrieval occurs sequentially, at least in the context of multisyllabic words (Dell et al., 1993; Houghton, 1990; Vousden et al., 2000). This view is certainly consistent with evidence from instances where phonological retrieval fails. For example, in the tip of the tongue state, the early phonemes of a word are more likely to be retrieved than later ones (Brown, 1991).

Similarly, in fluent aphasia, phonemic paraphasic errors are more likely to preserve the target's beginning phonemes than its later ones (Wilshire, 2002).

Of course, since our form-related distractors shared entire syllables with their targets, the results are also compatible with a view that syllables are retrieved in sequence, but their individuals segments are retrieved in parallel. The next step in this research will most likely involve examining different types of sharing (whole vs. part syllables), and perhaps a wider range of SOAs. The use of converging approaches may provide further insights– for example, the study of aphasic phonological errors and how they are influenced by form-related distractors, or the study the temporal and spatial signatures of form facilitation effects using EEG or MEG.

Footnotes

1. These 192 distractors had four to eight phonemes (mean 4.81), with Celex lexeme frequencies between 0 and 292 tokens per million (geometric mean 4.40). In 19 instances, where there were no appropriate distractor words, a best match was chosen that either shared all phonemes in the first syllable *except* the last one (e.g., *finger-finish*, 13 instances), or alternatively, additionally shared the first phoneme of the second syllable (e.g., *arrow-arab*, 6 instances).

2. These 192 distractors had three to eight phonemes (mean 4.84), with Celex lexeme frequencies between 0 and 428 (geometric mean 4.82). The mean time to overlap between each end related distractor and its corresponding target was 335ms (sd=97). In eight instances where there were no appropriate distractor words, the distractor shared all phonemes in the second syllable except the first one (e.g., *balloon-typhoon*).

3. The assignment of distractor alternatives to SOAs varied across targets. So for example, "poppy" was used as a begin-related distractor for *pocket* at -200ms and +400ms, but as an unrelated distract for *tiger* at +200ms and +400ms.

References

Abel, S., Dressel, K., Bitzer, R., Kümmerer, D., Mader, I., Weiller, C., & Huber, W. (2009). The separation of processing stages in a lexical interference fMRI-paradigm. *Neuroimage, 44(3),* 1113-1124. doi: 10.1016/j.neuroimage.2008.10.018

Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience, 23*(6), 1358-1367. doi:10.1162/jocn.2010.21519.

Brown, A. S. (1991). A review of the tip-of-the-tongue state. *Psychological Bulletin, 109*, 339–347. doi: 10.1037/0033-2909.109.2.204.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J (1993). PsyScope: An interactive graphics system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers, 25*(2), 257-271. doi: 10.3758/BF03204507.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Damian, M. F. (2003). Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29(3)*, 416-431. doi: 10.1037/0278-7393.29.3.416.

Damian, M. F., Bowers, J. S., Stadthagen-Gonzalez, H., & Spalek, K. (2010). Does word length affect speech onset latencies when producing single words? *Journal of Experimental Psychology: Learning, Memory and Cognition, 36(4)*, 892-905. doi: 10.1037/a0019446.

Damian, M. F., & Dumay, N. (2009). Exploring phonological encoding through repeated segments. *Language and Cognitive Processes, 24(5)*, 685-712. doi: 10.1080/01690960802351260.

Damian, M. F., & Martin R.C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 345–361. doi: 10.1037/0278-7393.25.2.345.

de Zubicaray, G. I. & McMahon, K. L. (2009). Auditory context effects in picture naming investigated with event-related fMRI. *Cognitive, Affective & Behavioral Neuroscience, 9*(3), 260-269. doi: 10.3758/CABN.9.3.260.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321. doi: 10.1037/0033-295X.93.3.283

Dell, G. S. (1988). The retrieval of phonological forms in production: Test of predictions from a connectionist model. *Journal of Memory and Language, 27*, 124–142. doi: 10.1016/0749-596X(88)90070-8.

Dell, G. S., Juliano, C., & Govindjee, A. (1993). Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science, 17*, 149–195. doi: 10.1207/s15516709cog1702_1

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43*, 182–216. doi: 10.1006/jmla.2000.2716

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*(1), 1-31. doi: 10.1006/jmla.1996.0001,

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 287–318). London: Academic Press.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92(1)*, 101-144. doi: 10.1016/j.cognition.2002.06.001

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition 24*(5), 1256–1274. doi: 10.1037/0278-7393.24.5.1256.

Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonologically related distractors in picture–word interference. *The Quarterly Journal of Experimental Psychology, 54A*(2), 371–382. doi: 10.1080/02724980042000273.

Johnson, V. E. (2005). Bayes factors based on test statistics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(5)*, 689-701. soi: 10.1111/j.1467-9868.2005.00521.x

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. doi: 10.1017/S0140525X99001776.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1–36.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1), 71-102. doi: 10.1016/0010-0277(87)90005-9.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1– 86. doi: 10.1016/0010-0285(86)90015-0.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effect of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 1146–1160. doi: 10.1037/0278-7393.17.6.1146.

Meyer, A. S., & van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review, 7(2),* 314-319. doi: 10.3758/BF03212987.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 5*2, 189 –234. doi: 10.1016/0010-0277(94)90043-4.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460–499. doi: 10.1037/0033-295X.107.3.460.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*, 249–284. doi: 10.1016/S0010-0277(97)00027-9.

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and readings their names. *Memory & Cognition, 32*(2), 212-222. doi: 10.3758/BF03196853.

Roelofs, A., Meyer, A. S., & Levelt, W. J. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(1)*, 246-251. doi: 10.1037/0278-7393.22.1.246.

Schriefers, H. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology, 11(1)*, 17-50. doi:10.1080/713752301

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*, 86–102. doi: 10.1016/0749-596X(90)90011-N

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, *132(12)*, 3411-3427. doi: 10.1093/brain/awp284.

Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition, 53*(2), 91-127. doi: 10.1016/0010-0277(94)90067-1.

Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition, 42*, 213–259.doi: 10.1016/0010-0277(92)90044-I

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language, 42*, 497-425. doi: 10.1006/jmla.1999.2693.

Strijkers, K., & Costa, A. (2011). Riding the lexical speedway: a critical review on the time course of lexical selection in speech production. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00356.

Wilshire, C. E. (2002). Where do aphasic phonological errors come from? Evidence from phoneme movement errors in picture naming. *Aphasiology, 16(1-2)*, 169-197. doi: 10.1080/02687040143000528.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology, 41*(2),101-175. doi: 10.1006/cogp.2000.0739.

Author Note

Carolyn Wilshire and Sunita Singh, School of Psychology, Victoria University of
Wellington, New Zealand and Catherine Tattersall, School of Human Communication
Sciences, University of Sheffield, United Kingdom.

Table 1

Mean naming latencies for correct responses (least squares means), collapsing across both subjects and items, with associated 95% confidence intervals in parentheses. To allow for comparison with other studies, estimates for Cohen's d_{av} have been calculated by dividing each mean difference by the average subjectwise standard deviation for the two conditions being compared, using the formula from Cumming (2012).

Condition	Target SOA			
	-200ms	0ms	+200ms	+400ms
Begin related	709 *(668-750)*	676 *(635-717)*	678 *(637-719)*	708 *(668-749)*
Begin unrelated	731 *(690-772)*	740 *(700-781)*	730 *(689-771)*	710 *(669-751)*
Mean difference	**22 *(6, 39)***	**64 *(48, 81)***	**52 *(35, 69)***	**2 *(-15, 18)***
Cohen's d_{av} (estimated)	**0.22**	**0.64**	**0.51**	**0.01**
End related	689 *(648-729)*	693 *(652-733)*	713 *(673-754)*	721 *(680-762)*
End unrelated	694 *(653-735)*	714 *(673-755)*	724 *(684-765)*	745 *(704-785)*
Mean difference	**5 *(-12, 22)***	**21 *(5, 38)***	**11 *(-8, 28)***	**24 *(7, 41)***
Cohen's d_{av} (estimated)	**0.06**	**0.26**	**0.13**	**0.24**

Figure Captions

Figure 1
Schematic illustration showing the various distractor onset timings used in the experiment, and how they might correspond to four major planning stages involved in target word production (all processes succeeding phonological retrieval are referred to collectively as "articulatory-motor processing"). The approximate onset and duration of each planning stage has been estimated based on the conclusions of Indefrey (2001) and Strijkers and Costa (2012), but also taking into account the greater delays likely imposed by the presence of auditory distractors, and by the use of bisyllabic targets. The diagram also illustrates how distractor onset was timed; that is, with respect to the onset of the crucial overlapping syllable. The critical hypotheses concerned the latest, 400ms distractor onset time, shown within a black box. At this time, lexical selection is likely to be fully complete and phonological retrieval partially complete. If all phonemes are retrieved in parallel, then begin-related and end-related distractors should be equally faciltatory when presented at this time. In contrast, if early phonemes are retrieved before later ones, then end-related distractors should elicit greater facilitation.
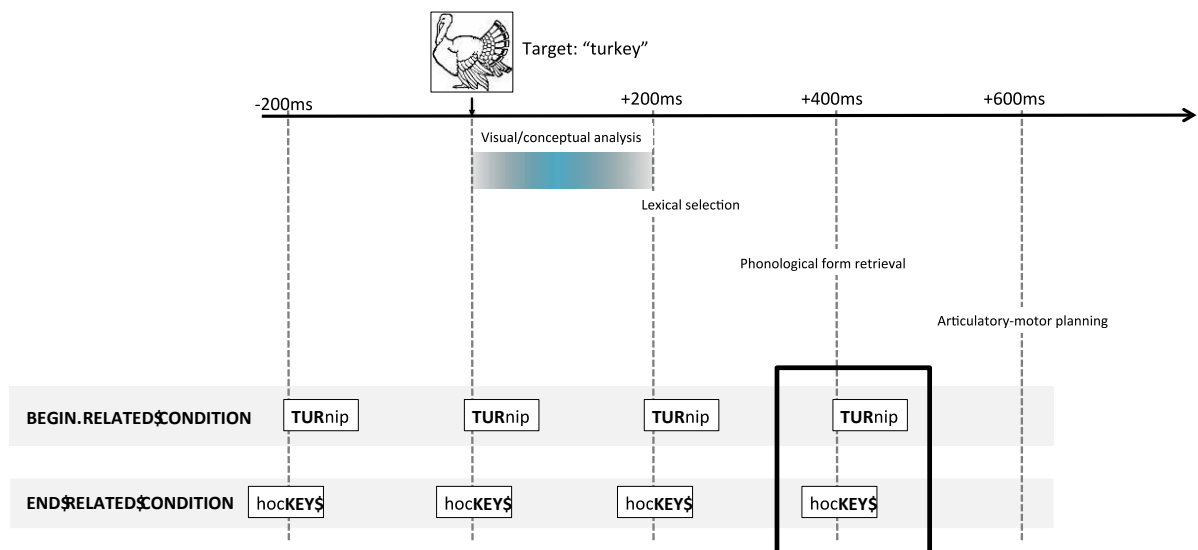
Figure 2
The mean latency difference between each phonologically related condition and its respective control for each of the four SOAs (derived from the least squares means). Error bars show 95% confidence intervals.