

# Effects of delay, length, and frequency on onset RTs and word durations

Romani, Cristina; Silverstein, Priya; Ramoo, Dinesh; Olson, Andrew

DOI:

[10.1080/02643294.2022.2070425](https://doi.org/10.1080/02643294.2022.2070425)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Romani, C, Silverstein, P, Ramoo, D & Olson, A 2022, 'Effects of delay, length, and frequency on onset RTs and word durations: Articulatory planning uses flexible units but cannot be prepared', *Cognitive Neuropsychology*, vol. 39, no. 3-4, pp. 170-195. <https://doi.org/10.1080/02643294.2022.2070425>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



# Effects of delay, length, and frequency on onset RTs and word durations: Articulatory planning uses flexible units but cannot be prepared

Cristina Romani, Priya Silverstein, Dinesh Ramoo & Andrew Olson

**To cite this article:** Cristina Romani, Priya Silverstein, Dinesh Ramoo & Andrew Olson (2022) Effects of delay, length, and frequency on onset RTs and word durations: Articulatory planning uses flexible units but cannot be prepared, Cognitive Neuropsychology, 39:3-4, 170-195, DOI: [10.1080/02643294.2022.2070425](https://doi.org/10.1080/02643294.2022.2070425)

**To link to this article:** <https://doi.org/10.1080/02643294.2022.2070425>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 19 Jun 2022.



[Submit your article to this journal](#)



Article views: 815



[View related articles](#)




[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

## Effects of delay, length, and frequency on onset RTs and word durations: Articulatory planning uses flexible units but cannot be prepared

Cristina Romani <sup>a</sup>, Priya Silverstein<sup>b</sup>, Dinesh Ramoo<sup>c</sup> and Andrew Olson<sup>c</sup>

<sup>a</sup>Department of Psychology, College of Health and Life Sciences, Aston University, Birmingham, UK; <sup>b</sup>Department of Psychology, Aston University, Birmingham, UK; <sup>c</sup>Department of Psychology, University of Birmingham, Birmingham, UK

### ABSTRACT

There is debate regarding whether most articulatory planning occurs offline (rather than online) and whether the products of off-line processing are stored in a separate articulatory buffer until a large enough chunk is ready for production. This hypothesis predicts that delayed naming conditions should reduce not only onset RTs but also word durations because articulatory plans will be buffered and kept ready. We have tested this hypothesis with young control speakers, an aphasic speaker, and an age and education-matched speaker, using repetition, reading and picture-naming tasks. Contrary to the off-line hypothesis, delayed conditions strongly reduced onset RTs, but had no benefit for word durations. In fact, we found small effects in the opposite direction. Moreover, frequency and imageability affected word durations even in delayed conditions, consistent with articulatory processing continuing on-line. The same pattern of results was found in CS and in control participants, strengthening confidence in our results. There is debate regarding whether most articulatory planning occurs offline (rather than online) and whether the results of off-line processing are stored in a separate articulatory buffer until a large enough chunk is ready for production. This hypothesis predicts that delayed naming conditions should reduce not only onset RTs but also word durations because articulatory plans will be buffered and kept ready. We have tested young control speakers, an aphasic speaker, and an age and education matched speaker, using repetition, reading and picture naming tasks. Contrary to the off-line hypothesis, delayed conditions strongly reduced onset RTs, but had no benefit for word durations. In fact, we found small effects in the opposite direction. Moreover, frequency and imageability affected word durations even in delayed conditions, consistent with articulatory processing continuing on-line. The same pattern of results was found in CS and in control participants, strengthening confidence in our results.

### ARTICLE HISTORY

Received 6 March 2021  
Revised 8 April 2022  
Accepted 10 April 2022

### KEYWORDS

Word production; aphasia; articulatory buffer; word durations; effects of preparation; word frequency; word length

In this study we ask if the human speech architecture includes an articulatory buffer where the articulatory commands for speech can be prepared ahead of time and stored or, alternatively, if memory resources required for articulation are encapsulated in the articulatory/motor planning system and can only be used during the actual production of speech.

There is a general agreement that word production requires a number of staggered stages involving: 1) Access to semantic representations; 2) Access to corresponding lexical representations as a sequence of symbolic units corresponding to phonemes; 3) Encoding of phonological representations for production; and 4) Articulatory/motor planning, where unpacked

phonological representations are converted into integrated articulatory gestures (Code, 1998; Goldrick & Rapp, 2007). Some authors identify a further stage of motor execution where sets of relevant muscles are activated with the appropriate synchrony and to the appropriate extent (see motor programming stage in the model FLF model by Van Der Merwe, 2009; 2021).

There is also a general agreement that working memory resources are needed during production processes. These resources are conceptualized as output buffers where linguistic representations are kept active while further processing is carried out. We assume that phonological encoding involves

**CONTACT** Cristina Romani  C.Romani@Aston.ac.uk

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

unpacking phonemes into features corresponding to articulatory targets (Kohn, 1984, 1989; Postma, 2000; see also articulatory phonology, Browman & Goldstein, 1992). This distinguishes phonological encoding from both the preceding stage of lexical access, where words are represented as sequences of unitary phonemes, and from the following stage of articulatory planning, where articulatory targets are converted into integrated articulatory gestures, i.e., integrated actions which specify how sequences of targets are reached. A buffer is needed to store the products of phonological encoding so that the planner has a look-ahead window to make the non-sequential adjustments necessary to realize the coordinated gestures needed for production. Movements need to be planned in synchrony with previous and following targets. Thus, even producing a single word will involve a phonological output buffer, although this will be minimally taxing (see Romani et al., 2011). This buffer will be taxed more by the need to retain multiple words in connected speech so that sentences can be assigned the proper prosodic contour. We will call the component which stores the products of phonological encoding the phonological output buffer, because it represents words in terms of phonological features, but it may be equally called a phonetic buffer if one wants to stress that features are articulatory rather than acoustic.<sup>1</sup>

The articulatory planner will convert phonetic representations into integrated gestures on-line, but it may also retrieve some gestures, corresponding to syllables or even larger chunks, in pre-packaged form from an articulatory store (for effects of syllable frequency in word production, see Laganaro & Alario, 2006; Levelt & Wheeldon, 1994 but also Croot et al., 2017 for no effects; for larger chunks see Varley & Whiteside, 2001). Depending on the model assumed, an articulatory buffer may also be needed to accrue and sequence these units before articulation starts. Some models assume that an articulatory plan corresponding to at least a word (but also possibly longer) needs to be ready before articulation can start (Klapp, 1995, 2003; Levelt et al., 1999; Meyer & Schriefers, 1991; Roelofs, 2002a; Wheeldon & Lahiri, 1997). These models must assume a buffer with an appropriate capacity (Klapp, 2003; Maas et al., 2008; Maas & Mailend, 2012; see later for details).

Models which assume that most of articulatory planning occurs online, instead, do not have the same need for an articulatory buffer. Information about the sequence of units to be produced would be held in the phonological buffer and only the units that are actually being converted would need to be kept active during production (minimalistic view of articulation; see Dell et al., 1993; Jordan, 1990; Kawamoto et al., 1999; MacKay, 1987; Santiago et al., 2000). Memory resources would be needed to guarantee that transitions occur smoothly without discontinuities and with the proper articulatory adjustments (e.g., on the basis of feedback and feed-forward information, see Hickok, 2012 for a review), but these memory resources would be limited and encapsulated in the process of articulatory realization (not available before that). We will reserve the term “buffer” for a component which, like the phonological buffer, accrues units before further processing (actual articulation) is initiated, keeps them in the right order, has a relatively large capacity, and can be refreshed, making representations available outside the process of articulation. We will contrast this with an online alternative where memory resources only involve temporary activation of representations.

Our study investigates the existence of an articulatory buffer and the related question of how much processing occurs after initiation of articulation. We compare effects of preparation, measured in terms of facilitation in delayed naming conditions, for onset RTs and word durations. Onset RTs measure the time from presentation of a stimulus to the beginning of articulation. They reflect the time taken for lexical access and, in most models, for phonological encoding to be completed (e.g., Buz & Jaeger, 2016; Kello, 2004). They may also reflect articulatory planning. When these processes can be prepared, RTs will be faster. Word durations measure the time from the beginning to the end of articulation. They only reflect articulatory processes which are carried out online (Buz & Jaeger, 2016; Meyer et al., 2007; Wheeldon & Lahiri, 1997). If articulatory plans can be made ready in an articulatory buffer, in delayed conditions word durations will be shortened because most of the planning will have been carried out beforehand. Instead, if most articulatory planning occurs online, preparation will have no effect. Therefore, the hypothesis that most of articulatory planning occurs offline predicts that there will be benefits of

delayed conditions in both onset RTs and word durations. The hypothesis that most articulatory planning is carried out online and that the memory requirements are minimal predicts that delay will benefit onset RTs only.

In addition, as further evidence for off-line vs. on-line planning, we will consider interactions between effects of preparation and word characteristics including length, frequency and imageability, both in onset RTs and word durations. Both the hypotheses at hand may predict an effect of length on onset RTs. Longer words will take longer to plan, if planning occurs offline; however, longer words will also take longer to encode if phonological encoding occurs in a sequential manner. Therefore, a length effect in onset RTs is consistent with both hypotheses. No length effect in onset RT, however, would only be consistent with the hypothesis that articulatory planning occurs online. If plans were prepared, longer words will need more time before speech is initiated.

Additionally, both hypotheses clearly predict effects of length on durations since longer words will take longer to say. However, if articulation can be prepared, the part of the length effect which is due to articulatory planning should reduce with preparation. Thus, the length effect on durations should weaken with delay. Instead, if articulatory plans cannot be prepared and buffered, the length effect on durations should be the same with or without delay. Effects of imageability and frequency are common in naming RTs (de Groot, *n.d.*; Perret & Bonin, 2019; Schwanenflugel & Stowe, 1989).<sup>2</sup> These effects could be linked both to the fact that high-frequency, high-imageability words have easier lexical access and to their having more practised articulatory programmes. Using the same logic applied to the length effect, if preparation of plans is possible, imageability and frequency should only influence onset RTs. Instead, if no preparation is possible, they may also influence word durations.

Before moving to describe, in detail, our experimental investigation, we will briefly review what is known about effects of preparation on word RTs and possible interactions with length, as well as evidence that word durations are sensitive to psycholinguistic effects. There are no studies that we know of that have measured effects of preparation on word durations.

## **Effects on onset RTs**

### **Effects of preparation and sequence length on onset RTs**

A number of studies have considered effects of preparation and length on onset RTs when participants produce sequences of syllables (Klapp, 2003; Klapp et al., 1973) or words (see Sternberg et al., 1978). Results were variable. Klapp (2003; Klapp et al., 1973) found a length effect in immediate production conditions which disappeared with preparation, but also the opposite pattern: no length effect in immediate conditions, but a length effect after preparation. He attributed these differences to whether or not participants integrated the syllables into a single articulatory plan.

Other studies with control and aphasic participants with apraxia of speech (AoS) have also produced contradictory results. Deger and Ziegler (2002) found a length effect with preparation in the control group, but not in the AoS group, while Maas and Mailend (2012) found the opposite. All authors assumed that people with AoS failed to integrate syllables into a single motor plan. However, Deger and Ziegler (2002) assumed that integration normally takes more time (despite preparation), producing a length effect in the control speakers, but not in the apraxic speakers, while Maas and Mailend (2012) assumed integrated programmes take less time to initiate, resulting in a lack of length effect in control speakers. Both assumptions are plausible: longer plans will take more time to integrate, but once integrated, plans will take less time to initiate. Without knowledge of the time course of integration, however, any set of results is compatible with any hypothesis. Both the inconsistency of the results and the fact that these paradigms ask for the production of novel syllable sequences, rather than familiar words, cast doubt on the locus of these effects. They could arise in encoding, rehearsal or retrieval of syllables rather than in articulatory programming. Experiments assessing effects of length with real words could be more revealing, but here, effects of length have been, at best, weak and inconsistent.

### **Effects of word length on onset RTs**

In reading, effects of length are limited to certain conditions. They are demonstrated only in non-proficient readers (e.g., Mason, 1978), and/or with low frequency

words (Weekes, 1997 see also Balota et al., 2004; Yap & Balota, 2009 for larger effects with low frequency words). In children, the length effect decreases with reading proficiency (Marinelli et al., 2016; Spinelli et al., 2005) and disappears with age in an orthographically opaque language like English (Marinelli et al., 2016). Moreover, effects of length have been found in tasks, like lexical decision, which do not involve spoken production (for a review see Barton et al., 2014), suggesting that any effect is due to orthographic decoding rather than to articulatory preparation. In picture naming, where serial orthographic decoding is not an issue, results have been inconsistent. Some studies have shown a length effect (Roelofs, 2002b), but others have shown an effect only in certain conditions (e.g., where stimuli are presented blocked by length; Meyer et al., 2003, 2007), or no effect at all (Bachoud-Lévi et al., 1998; Santiago et al., 2000, 2002), and some positive results have not been replicated with more controlled sets of stimuli (Damian et al., 2010). These inconsistent results do not support off-line planning. Assessing length effects in people with AoS could be more revealing, since effects could be magnified by a selective impairment in speech production, but we do not know of any existing studies.

### ***Effects of frequency and imageability on word durations***

In our experimental investigation, we will compare effects of preparation on onset RTs and word durations. Therefore, it is important to ascertain that word durations are sensitive to psycholinguistic variables.

Word durations are likely to be impacted in different ways by different factors. In connected speech, words are produced with longer durations when the message is ambiguous and/or when words have more neighbours (see; Gahl et al., 2012). Instead, they are produced with shorter durations in predictable contexts (Arnold, 2016), when words are repeated (Shields & Balota, 1991), and when they have higher frequency (Gahl, 2008; Lohmann, 2018; but also see Clopper for variability depending on conditions). Similarly, in delayed naming, word durations are longer for phonologically similar words (Mooshammer et al., 2009; Buxo-Lugo et al., 2020), for words with more phonological neighbours (Buz

& Jaeger, 2016), and for similar compared to identical pairs for words (e.g., tape-tape vs. tape-cape or tape-take; Mooshammer et al., 2009). Instead, durations are shorter for repeated syllables (Lam & Watson, 2010, 2014) and repeated words (Kahn & Arnold, 2015). These results show that a need for clarity, which arises with the presence of phonological neighbours or with an ambiguous context, will extend durations by encouraging full articulation of phonemes. Instead, articulatory practice—tapped through word repetition or word frequency—will make articulatory planning easier and reduce durations by increasing coarticulation between phonemes. Jacobs et al. (2015) have shown that words are produced with shorter durations after the overt production of an identical word, but not after silent reading. This is consistent with shorter durations reflecting a specific facilitation of articulatory planning (see also Goldrick et al., 2019).

Evidence that word frequency affects durations in naming is limited. Varley et al. (1999) only found a non-significant trend in control speakers. However, number of participants were very small ( $N=3$  control speaker, and 3 aphasic speakers and 4 with AoS). It has also been shown that in control speakers high-frequency words are produced with shorter onsets (Kawamoto et al., 1999) and shorter vowels (Munson, 2007), but analyses of whole words are lacking. Kello (2004) have shown that, in combination, stimulus variables including printed frequency, neighbourhood size, and regularity predict a significant amount of variance of word durations in a standard reading task. Moreover, Kello (2004) showed that in a tempo-naming task—where words had to be produced in synchrony with the beat of a metronome—the contribution of these variables increases with quicker tempos, suggesting more processing occurring during production. However, neither study analysed word frequency separately.

Finally, some studies have shown longer word durations in naming conditions with high semantic interference as in Stroop (Kello et al., 2000), continuous naming and cyclic-blocked naming tasks (Fink et al., 2018), with stronger effects when participants are under more time-pressure (Kello et al., 2000, 2004). Again, it is possible that with time-pressure, some of the processing normally occurring before articulation can be shifted online (e.g., suppressing competitors or monitoring for errors) slowing down production (but



for no effects see Damian, 2003; Goldrick et al., 2019). Regardless of the right explanation, these effects imply that linguistic processing may not be completed when articulation starts so that effects of psycholinguistic variables can influence durations. Instead, the hypothesis that all processes are completed beforehand predicts an impact only on RTs, with no impact on durations. In our experiments, we will test these hypotheses by considering effects of frequency and imageability in conditions of immediate vs. delayed naming. Both frequency and imageability may be a proxy for more practised articulatory programmes. However, effects of imageability on durations will be more indirect and depend more on whether any semantic effect can trickle down to affect articulatory planning stages.

### Plan of study

We will assess benefits of preparation for onset RTs and word durations by comparing delayed with immediate naming conditions and assessing interactions with length, frequency and imageability effects. These effects will be investigated in three production tasks: repetition, reading and naming. Although naming is more sensitive to lexical access than the other two tasks, all three tasks equally involve post-lexical stages of phonological encoding and articulatory planning. Therefore, any effects on word durations should be similar across tasks, effectively providing a replication of our results in different conditions.

The hypothesis that an articulatory plan cannot be prepared predicts no shortening of word duration in delayed response conditions. Null effects, however, could be due to lack of sensitivity when articulation is well-practised. To address this issue, tasks were administered not only to unimpaired, control speakers, but also to an aphasic participant with evidence of articulatory difficulties (a form of apraxia of speech, AoS).

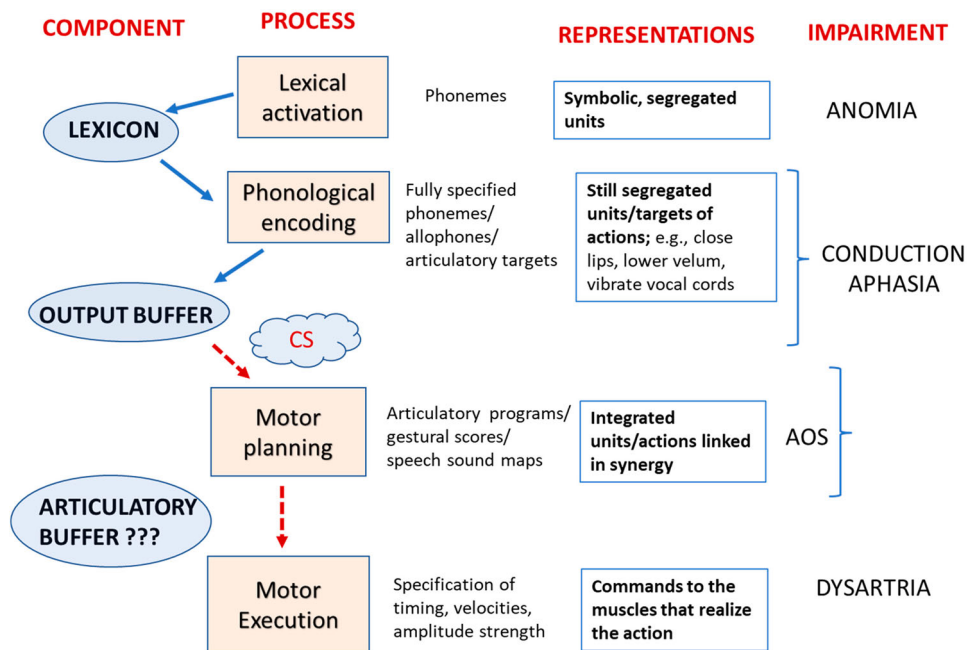
The case of CS has been extensively described in a single-case study (Ramoo et al., 2021). The characteristics of his speech production indicate that his main difficulty occurs after lexical access. CS, in fact, is impaired across production tasks (repetition, reading and naming) and, most importantly, makes similar types of errors and in similar proportions, across these tasks. The errors made by CS also reveal the nature of his speech impairment. First, he makes a high proportion of phonetic errors and syllabified

responses together with phonological errors. These are universally considered the hallmarks of an articulatory impairment (a type of apraxia of speech; Deger & Ziegler, 2002). On the basis of further analyses, we have argued, however, that CS's articulatory impairment is of a particular type and it involves timely feeding of phonological information to the articulatory planner.

Many of CS's errors involve repeated attempts at the target. Sometimes these are completed attempts, but more often they are false starts (which could be correct or incorrect) followed by more complete and fluent productions. This impairment, where articulatory plans are built slowly from progressively larger blocs, contrasts with other apraxic impairments where errors are phonological simplifications of the target, similar to those produced by children. These errors are likely to be motivated by an inability to compute gestures with complex spatio-temporal parameters which are beyond the articulatory competency of the speakers (Galluzzi et al., 2015; Romani et al., 2011, 2017; Romani & Galluzzi, 2005). CS also makes some of these errors, but in most cases his production difficulties (phonetic errors, syllabifications) are overcome by repeated attempts, suggesting he has limited capacity for transferring information to the planner, which can be overcome by working on smaller chunks of information at a time. This impairment makes CS the ideal candidate to examine whether articulatory plans can be prepared and buffered. If more preparation time is useful for articulatory planning, it should be especially useful in CS. Figure 1 shows a schematic model of speech production with the locus of CS' impairment.

Our predictions are summarized below (see Table 1). The hypothesis that an articulatory plan must be prepared offline and stored in an articulatory buffer before articulation starts (possibly holding a sequence of articulatory gestures corresponding to the phonological word) predicts that:

- (1) Preparation will reduce word durations because a more complete plan will be ready at the "go" signal and limited planning will have to be done after speech starts. Moreover, preparation will reduce dysfluencies in CS, who has trouble compiling whole-word programmes and uses a *conduite d'approche*, where progressively longer



**Figure 1.** Schematic model of speech production. Alternative terminology is shown for the representations involved at different stages to allow integration with current literature. The model shows hypothesized deficit of CS in terms of slow transfer on information and consequences for further stages. These consequences should be ameliorated with preparation if articulatory programmes can be buffered.

**Table 1.** Predictions models where articulation can be prepared and buffered as phonology or where articulation cannot be buffered before production.

	ARTICULATION IS OFFLINE & PREPARED IN A BUFFER					ARTICULATION IS ONLINE & NOT PREPARED LIKE PHONOLOGY				
	Onset RTs		Word durations		Errors	Onset RTs		Word durations		Errors
	Controls	CS	Controls	CS	CS	Controls	CS	Controls	CS	CS
<b>SIGNIFICANT EFFECT</b>										
<b>Delayed vs immediate naming</b>										
Repetition	YES: faster	YES: faster	YES: shorter	YES: shorter	Fewer	YES: faster	YES: faster	NO: no change	NO: no change	no change
Reading										
Naming										
<b>High frequ/imm. better than low</b>										
Repetition	YES: faster	YES: faster	No: same	No: same	fewer	YES: faster	YES: faster	YES: shorter	YES stronger than SS	fewer
Reading										
Naming										
<b>Frequ./imm effect weaker in delayed conditions</b>										
Repetition	YES	YES	YES	YES		YES	YES	NO	NO	
Reading										
Naming										
<b>Longer words worse than shorter</b>										
Repetition	YES: slower	YES: slower	YES: longer	YES: longer		Unclear	Unclear	YES: longer	YES: longer	
Reading										
Naming										
<b>Length effect weaker in delayed conditions</b>										
Repetition	YES	YES	YES	YES		YES	YES	NO	NO	
Reading										
Naming										

Cells in different colours contrast different predictions for different models.

chunks of the word are produced from smaller chunks.

(2) Effects of frequency/imageability will be present in onset RTs, but not in word durations. This is because if planning is completed before

articulation starts, these variables should no longer affect durations.

(3) Effects of word length should be present because longer words will take longer to encode and to plan for articulation.



- (4) Obviously, longer words will take longer to say, but the length effect should be reduced in delayed naming because planning would occur beforehand so that only time of execution will affect durations.

Preparation effects and interactions with length and frequency may be particularly strong in CS who has dysfluent speech.<sup>3</sup>

In contrast, the hypothesis that articulatory planning occurs mostly online and articulatory plans cannot be prepared and buffered predicts that:

- (1) The opportunity to prepare will strongly reduce onset RTs, but not word durations both in control speakers and in CS. Similarly, CS's errors and repeated attempts will not be reduced in delayed compared to immediate naming conditions.
- (2) Effects of word frequency/imageability will be present in both onset RTs and word durations since articulatory compiling occurs online and it will be faster for more practised programmes.
- (3) Effects of word length may or may not be present in onset RTs if phonological depending on whether phonological encoding occurs serially or in parallel (e.g., see Roelofs, 2002b)
- (4) Any effect of frequency/imageability/length on word durations should not disappear or weaken in delayed conditions since articulatory programmes cannot be prepared.
- (5) The same types of effect will be present in in CS who has dysfluent speech and in control speakers.

In addition, we expect that, if our proposed locus of impairment is correct, CS will: a) make a high rate of repeated attempts in all three production tasks; b) make more repeated attempts on longer words where more information needs to be transferred to articulatory planning; c) produce longer durations, even on correct items, than an age- and education-matched control. Finally, if results from CS can strengthen the interpretation of results from neurotypical speakers, controls' results can also strengthen our interpretation of CS's impairment. If CS shows no effect of preparation, this could happen because the capacity of an articulatory buffer is so reduced that it does not allow any storage. However, this hypothesis is unlikely if similar effects are shown by unimpaired speakers where all speech components are assumed to be intact.

We will first summarize CS's case report. We will then present method and results for computerized tasks administered to CS, a matched control, and a group of younger control speakers.

### CS Case study

At the time of testing, CS was a 75-year-old, right-handed man who had suffered an ischaemic stroke two years prior to testing. His CT scan showed a wedge-shaped area of low attenuation in the left parietal region (middle cerebral artery territory) with some normal density within it. This indicates partial infarction with some tissue perfusion in the damaged area. CS had a B.Sc. (Hons) in Electrical Engineering from Aston University and had worked as an engineer for the BBC before taking early retirement. At the time of testing he was married and enjoyed an active life, with lots of hobbies. He liked sports and had previously played hockey, golf and badminton. After his stroke he engaged in orienteering, hill walking and working in his allotment. He was recruited for a research study via the South Birmingham Community Support Centre of the Stroke Association and was tested at the University of Birmingham between 2012 and 2014. A detailed case study of CS is described in Ramoo et al. (2021). We refer readers to this paper for additional details and only summarize here the main features of his performance necessary to motivate the present investigation.

CS's speech was grammatical with a good range of words but a halting quality, characterized by false starts, syllabified words, phonetic and phonological errors. CS's speech was also characterized by a marked *conduite d'approche* where target words were built up from progressively larger speech units. These repeated attempts (RAs) were often followed by a correct and fluent response. A general investigation of CS's language abilities revealed good phonological input processing, good semantic and lexical processing, good sentence comprehension and good phonological short-term memory (good digit span, probe rhyme span and probe semantic span). Instead, word production was impaired across tasks (repetition, reading, and naming) with a similar prevalence of phonological errors (mainly non-lexical) with similar characteristics, consistent with a post-lexical deficit (Goldrick & Rapp, 2007). Difficulties in lexical access would predict normal or much better

performance in word repetition (where the stimulus only needs to be reproduced) which was not the case.

Our experimental investigation confirmed what was observed clinically. He made high rates of phonetic errors (repetition: 20.2%; reading: 14.4% of stimuli) and syllabifications (repetition: 22.4%; reading: 15.1% of stimuli) and a high number of repeated attempts in repetition and reading (9% of stimuli in both tasks). RAs occurred in two forms. Sometimes CS produced a complete, but erroneous response that he tried to revise, being successful about half of the time (e.g., algebra /ældʒɪbrə/ > ældʒɪblə ... ældʒɪbrə; flax /flæks/ > flæfjs ... flæks). Other times he produced an incomplete response (i.e., a fragment), correct or incorrect, followed by more complete attempts which generally led to a successful outcome (e.g., hospital > hɒf ... hɒspɪtəl; inhibition > ɪnhɪ ... ɪnhɪ ... ɪnhɪbj ... ɪnhɪbjən). Taken together these characteristics are similar to a *conduit d'approche* where a correct response is built up through repeated attempts, guided by an intact lexical representation.

CS's phoneme errors were affected by word position, progressively increasing towards the end of words, but not affected by word length independent of this positional effect. This is consistent with a deficit feeding phonological information to the articulatory planner. Delays accumulate across the word, causing increasing difficulty computing the right articulatory plan. This pattern is less consistent with a classical buffer impairment where one would expect word length to have an independent effect (with difficulties influenced by the overall number of phonemes that must be stored in the buffer). Finally, CS showed significant effects of word frequency, with more errors on low frequency words in all three tasks (repetition, reading, and naming), and effects of phonological complexity, with more errors on complex phonemes and consonant clusters.

Taken together, CS's speech characteristics are well explained by noise and delays in accessing/compiling articulatory plans, which are worse when computing gestures for more complex phonemes and/or for less practised/low frequency words. This problem would also explain the other salient characteristics of his speech. Syllabifications would result directly from dysfluencies in receiving information, phonetic errors from receiving noisy or competing information and RAs from retrying production to resolve errors and dysfluencies. Note that, given CS's phonological

errors and *conduit d'approche*, he might have been classified as having conduction aphasia rather than apraxia of speech. However, a problem in feeding information to the articulatory planner provides a single explanation for all of his speech characteristics.

In our experimental investigation we will compare CS to a right-handed control speaker matched for age and education (his wife: SS; 75 years old with a Bachelor of Arts degree). Our main prediction in comparing CS and SS is that CS will show significantly longer word durations across tasks. This will confirm his lack of fluency. In addition, it is likely that onset RTs will be slower in CS. Speed of processing is generally reduced after brain damage and minor lexical impairments are very common. Therefore, it is likely that effects of frequency and length will be stronger in CS, both in terms of RTs and accuracy. Equally, it is likely that he will show stronger advantages of delay which should reduce the difference in onset RTs compared to SS. Crucially, the hypothesis that articulatory representations cannot be buffered makes the strong prediction that CS's fluency will not improve in delayed conditions either in terms of word durations or repeated attempts. If control participants also show no benefits of delay, this will be convergent evidence from both neuropsychological and control data that articulatory representations cannot be prepared and stored independent of the process of speaking.

## Experimental investigation

### Method

#### Neurotypical participants

In addition to CS and SS, tasks were administered to 18 younger typical speakers (average age = 22.6, SD = 2.64; Female/male = 10/8). The younger controls were all students (9 undergraduates, 9 postgraduates) who completed the experiment to obtain research credits. 17/18 were right-handed. Some results from the younger participants were excluded due to technical errors (the reading length subtask for one participant and the naming length subtask for 2 participants).

#### Procedure

Each task was presented using EPrime (Psychology Software Tools, Pittsburgh, PA). Participants were asked to repeat, read, or name the stimuli as quickly

and as accurately as possible. Responses were recorded using an Audio Technica AT8035 microphone and a TASCAM digital recorder. A XENYX Q802USB mixer was used to simultaneously send the audio signal to the digital recorder and to a Cedrus SV-1 voice key to measure onset RTs.

Stimuli from different categories were presented in a random order. The participant was seated approximately 60 cm from the computer screen, so that stimuli subtended roughly nine degrees of visual angle. Before the presentation of each stimulus, a cross appeared at the centre of the screen for 500 ms. It was immediately followed by the stimulus: the spoken word for repetition, the written word for reading or the picture for naming. The task was different from a standard production task because the participants had to produce the response only when they were given a “go” signal after a variable delay. In the case of repetition, the “go” signal was the appearance of a green square (presented together with the start of the word in repetition, delay 0, to match reading and picture naming). In the case of reading, the word turned green; in the case of picture naming, a contour around the picture turned green. The different delays were intermixed during the task (for a schematic outline of the procedure see [Figure 2](#)).

The younger controls completed the tasks in two separate sessions at least one week apart, with all three tasks and a version of each list attempted in each session. CS and SS completed the tasks in four weekly sessions since we were more mindful of possible fatigue. Each session presented tasks separated by 15-20-min breaks.

### Materials

The same lists were administered in repetition and reading. Different lists were administered for picture naming since the range of words that can be elicited unambiguously with pictures is more restricted and generally excludes abstract/less frequent words. All stimuli were nouns. For repetition and reading we administered the following lists:

- A list assessing effects of frequency and imageability had 72 items (36 high frequency and 36 low frequency). In each frequency category, half of the items were high imageability and half were low imageability. High- and low-frequency words and

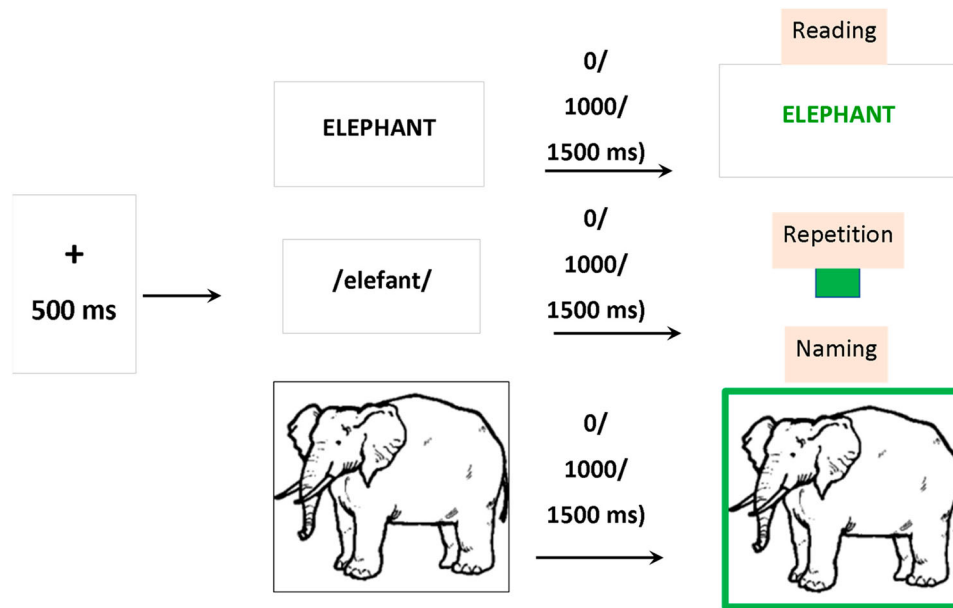
high and low-imageability words were matched for phoneme length (HF: length = 7.1, SD = 1.75; LF: length = 7.3, SD = 1.62; HI: length = 7.2, SD = 1.61; LI: length = 7.3, SD = 1.76).

- A list assessing effects of length had 36 items (18 short words, 4–6 phonemes long, and 18 long words, 7–9 phoneme long). Words of different lengths were matched for frequency (short words: log frequency = 2.1, SD = 1.2; long words: log frequency = 2.0, SD = 1.2).

For picture naming we administered the following lists:

- A list assessing effects of frequency had 48 items, half high frequency and half low frequency, matched for length (HF: length = 5.3, SD = 0.44; LF: length = 5.3, SD = 0.4)
- A list assessing effects of length had 72 items; 12 words for each of 6 lengths (3, 4, 5, 6, 7 and 8 phonemes) matched for frequency (3-phonemes: log frequency = 2.4, SD = 0.63; 4-phonemes: log frequency = 2.2, SD = 0.75; 5-phonemes: log frequency = 2.3, SD = 0.72; 6-phonemes: log frequency = 2.3, SD = 0.81; 7-phonemes: log frequency = 2.3, SD = 0.83; 8-phonemes: log frequency = 1.2, SD = 0.84).

Words in contrasting categories were also matched for syllabic complexity (by considering the number of complex clusters and hiatuses), and for the number of items starting with either vowels or fricatives. Stimuli were categorized as “low frequency” if they had a frequency of <150 tokens per million, and “high frequency” if they had a frequency of >1000 tokens per million (Celex database; Baayen et al., 1993). Imageability was assessed using the MRC Psycholinguistics Database (Coltheart, 1981) which specifies values between 100 and 700 ( $M = 450$ ,  $SD = 108$ ). Words with values >450 were considered high imageability and values <450 were considered low imageability. Since some low frequency words in our list did not have imageability values in the database, we asked a group of eight participants to rate imageability using a Likert scale from 1—to 10, with 10 corresponding to words that evoked a clear mental image. We considered words high imageability if they scored >5 and low imageability if they scored <5.



**Figure 2.** Schematic diagram of stages in the computerized tasks.

The go-signal to produce the words appeared after delays of 0, 1000, or 2000 ms for the reading and repetition lists and after delays of 0, 1000, 1500 or 2000 for the picture naming lists. For naming, we included an intermediate delay since naming RTs are longer than reading/repetition RTs. We did not want to miss a delay where preparation was useful (before a potential decay of buffered representations).

All lists were presented twice in separate sessions, associated with different delays in each version. If a word was presented with shortest delay in session one, it was presented with the longest delay in session 2, and *vice versa* with intermediate delays occurring randomly in the two sessions. Overall, 216 stimuli were presented for repetition and reading and 240 for naming.

The pictures were colour photographs approximately  $500 \times 500$  pixels. The words were written in capital letters, Arial font 16pt. The spoken words for repetition were recorded by a male native English speaker.

### Analyses

Errors were coded offline, blind to the category of the target. Word durations were computed by hand, using the acoustic analysis software Praat (Boersma, 2001). Errors (the wrong word for the target) were excluded from the RT and duration analyses. The percentage errors were, for younger controls: 2% in

repetition, 3% in reading, 15% in naming; for SS, 0.3% in repetition, 0.3% in reading, 13.8% in naming; for CS, 20% in repetition, 28% in reading, 34% in naming. For the RT and duration analyses, we excluded synonyms (e.g., “teacher” instead of “professor” or “laptop” instead of “computer”). These items were counted as “correct” in the accuracy analyses. Fourteen of CS responses were eliminated for this reason (5.8%) and between 1 and 11 percent of young control responses (mean = 7.5%). We also excluded responses that were initiated  $>3$  SD from the participant mean and/or responses that were initiated prior to 200 ms (anticipations). The percentage of trials that were outliers or anticipations were, for younger controls, 4% in repetition, 5% in reading and 4% in naming; for SS, 2% in repetition, 2% in reading, 3% in naming; for CS, 2% in repetition, 6% in reading and 3% in naming.

For each task and list, results were statistically analysed with generalized linear models (for CS and SS) or linear mixed models (for younger controls) using either onset RTs or word durations as the dependent measure and the psycholinguistic variables (imageability, frequency, length) and the condition (type of delay) as within-subject variables. The mixed models for younger controls included a random intercept for participants. The statistical significance of terms was evaluated by comparing models with and without the critical terms using likelihood ratios. Likelihood ratio chi-square values from model

comparisons are marked  $G^2$  to distinguish them from chi-square values from a traditional test of independence (Agresti, 2013, p. 76). Frequency and imageability were categorical variables. Length was a continuous variable. We will only report analyses which considered “delay” a categorical variable (0 vs. >0) since the main results were the same across delays >0. In comparing CS and SS, ‘participant’ was included as a between-subjects variable. We will only report results for two-way interactions which are relevant to our hypotheses. Error rates were analysed with binomial logistic regression (generalized linear models with a binomial link function).

## Results

### Accuracy

Table 2 shows error rates in different conditions. Only results for younger controls and CS are reported. SS made too few errors to analyse (no errors in repetition and reading; very few errors in picture naming: 8/144 = 5.6%, on the length list and 11/96 = 11.5% on the frequency and imageability list). The younger controls made more errors on low than high frequency words in reading ( $G^2(2) = 25.8, p < .001$ ) and naming ( $G^2(2) = 29.7, p < .001$ ), but not repetition ( $G^2(2) = 3.1, p = .21$ ). CS made significantly more errors on low than high frequency words in reading ( $G^2(1) = 5.4, p = .02$ ), but not in repetition ( $G^2(1) = 0.08, p = .78$ ) or naming ( $G^2(1) = 0.76, p = .38$ ). For CS, the relatively small number of errors limited our ability to detect differences. There were no significant effects of imageability. There were no effects of length in the younger controls. CS made more errors on longer words in naming ( $G^2(1) = 8.0, p = .005$ ) and when all three tasks were considered together ( $G^2(1) = 6.20, p = .01$ ).

Importantly, there were no benefits of delay on overall error rates.

Delay did not affect rate of errors either in the controls or in CS. If something went astray during lexical retrieval, it was not modulated by a short preparatory delay.

Turning to repeated attempts, CS made more repeated attempts on longer words in naming ( $G^2(1) = 5.9, p = .005$ ), and when all the three tasks were considered together ( $G^2(1) = 6.1, p < .01$ ; reading and repetition had non-significant downward trajectories with length). Delay did not reduce the number of repeated attempts (longer delays

increased the number of repeated attempts in repetition,  $G^2(1) = 3.78, p = 0.05$ ).

These results confirm with additional materials the impairment hypothesized for CS by Ramoo et al. (2021). CS shows an impairment across all production tasks. He shows slightly more errors in naming and reading than repetition, consistent with a mild deficit in lexical access. However, the type of errors made are similar across tasks with high rates of repeated attempts which increase on longer words. Together with high rates of phonetic errors and syllabifications, these results indicate a deficit in converting phonological representations into articulatory gestures. Importantly, however, repeated attempts do not reduce when CS has a chance to prepare, in contrast with the hypothesis that articulatory plans can be prepared and buffered.

### Onset RTs

Figure 3 shows results for onset RTs with Panel A showing effects of word frequency and imageability, and Panel B showing effects of length. Corresponding statistical analyses are reported in Tables 3 and 4. In these and following tables, to improve readability, we report results separately for delay = 0 and delay > 0 which makes interactions evident without listing them explicitly.

**Main effect of participant (CS vs SS).** CS was not systematically slower than SS. With the frequency/imageability list, he was, in fact, faster in reading ( $F(1) = 35.7, p < .001$ , marginal  $R^2 = 6.0$ ). With the length list, he was faster in reading ( $F(1) = 31.5, p < .001$ , marginal  $R^2 = 12.7$ ) and in naming ( $F(1) = 7.54, p = .007$ , marginal  $R^2 = 1.9$ ), and slower in repetition ( $F(1) = 13.1, p < .001$ , marginal  $R^2 = 8.8$ ). Effects of delay were stronger in CS than SS, in repetition and reading with both the frequency and the length lists. These results show that CS can take full advantage of preparation, at least when lexical access and phonological encoding are concerned.

**Effects of frequency/imageability.** Younger controls were slower on words that were both low frequency and low imageability when there was no delay in reading and repetition (but not naming). This is responsible for the main effects of imageability and frequency in reading and repetition in addition to the interaction. Younger controls also produced a



**Table 2.** % of errors in computerized tasks.<sup>4</sup>

	Younger controls				CS					
			G <sup>2</sup>				G <sup>2</sup>			
			N	p-value			N	p-value		
	FREQUENCY				FREQUENCY					
	<b>High</b>	<b>Low</b>			<b>High</b>	<b>Low</b>				
Repetition	1.6	2.2	2592	n.s.	15.3	19.4	144	n.s.		
Reading	<b>0.3</b>	<b>2.6</b>	2592	<.001	<b>8.3</b>	<b>41.7</b>	144	<b>.02</b>		
Naming	<b>6.4</b>	<b>16.8</b>	1632	<.001	25	39.6	96	n.s.		
	IMAGEABILITY				IMAGEABILITY					
	<b>High</b>	<b>Low</b>			<b>High</b>	<b>Low</b>				
Repetition	2.1	1.7	2592	n.s.	13.9	20.8	144	n.s.		
Reading	1.2	1.8	2592	n.s.	22.2	27.8	144	n.s.		
	LENGTH				LENGTH					
	<b>Short</b>	<b>Long</b>			<b>Short</b>	<b>Long</b>				
Repetition	0.3	1.4	1296	n.s.	22.2	22.2	72	n.s.		
Reading	0	1	1224*	n.s.	19.4	25	72	n.s.		
Naming	7.8	6.3	2304**	n.s.	<b>21.9</b>	<b>37.5</b>	144	<b>.005</b>		
	DELAY				DELAY					
	<b>0</b>	<b>1000</b>	<b>2000</b>		<b>0</b>	<b>1000</b>	<b>2000</b>			
		(1000/1500 naming)				(1000/1500 naming)				
Repetition	2	1.9	0.8	3888	n.s.	15.3	18.1	23.6	216	n.s.
Reading	1	1.4	1	3816*	n.s.	27.8	25	19.4	216	n.s.
Naming	7.8	11	6.5	3936**	n.s.	33.3	26.7	30	240	n.s.
	<b>REPEATED ATTEMPTS</b>				LENGTH					
					<b>Short</b>	<b>Long</b>				
Repetition					6.3	11.4	67	n.s.		
Reading					16.7	22.9	71	n.s.		
Naming					<b>21.3</b>	<b>31.9</b>	141	<b>.02</b>		
Total					<b>17.3</b>	<b>23.1</b>	279	<b>.04</b>		
	DELAY				DELAY					
	<b>0</b>	<b>1000</b>	<b>2000</b>		<b>0</b>	<b>1000</b>	<b>2000</b>			
Repetition					<b>4.4</b>	<b>11.4</b>	<b>14.3</b>	208	0.05	
Reading					27.5	22.7	20.7	193	n.s.	
Naming					28.6	24.1	31	230	n.s.	
Total					19.8	19.5	21.5	631	n.s.	

HF = high frequency words; LH = low frequency words. In naming, Short = words 4-5-6 phonemes; Long = words of 6-7-8 phonemes. In reading and repetition, Short = words 4-5-6 phonemes; Long = words of 7-8-9+ phonemes. Significant results are in bold. \* One participant excluded because of a technical recording error; \*\* two participants excluded because of a technical recording error. n.s. = non significant  $p$  value  $>.05$ ; Significant results are in bold.

paradoxical effect in naming with delay  $> 0$  (with low frequency words being faster see Table 3). CS showed stronger frequency effects than SS in reading and repetition after a delay. No other effects reached significance.

**Main effects of delay.** There were strong main effects of delay. A delay before the response significantly reduced onset RTs in all tasks, across all lists, and in all participants (Tables 3 and 4). Therefore, in all participants, the possibility to prepare speeded up the stages preceding articulatory initiation.

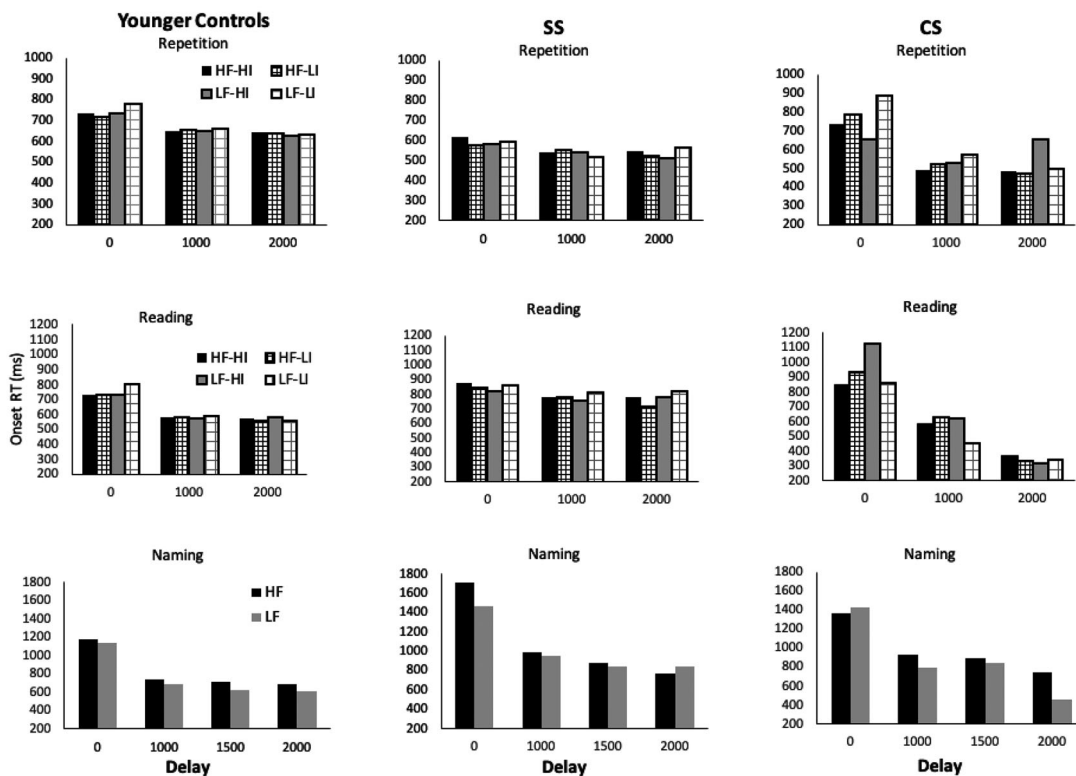
**Interactions: frequency/imageability X delay.** In the young controls, the effect of low frequency/low imageability words was significant only at zero delay. These results are expected. Lexical access may be particularly difficult for words that are low both in

imageability and frequency, but these effects can disappear when preparation is allowed.

**Length and length X delay.** The hypothesis that articulatory planning occurs offline predicts there should be a length effect in immediate naming conditions which weakens with the ability to prepare. Effects of length, instead, were weak and inconsistent. In naming, younger controls showed the expected length effect with no delay (longer words initiated later). This effect not only disappeared but unexpectedly reversed with delay. In reading and repetition, they showed no significant effects. In CS, there was a marginal length effect in naming at 0 delay and in reading at delay  $>0$ . In SS, there was a length effect only in repetition after a delay. Comparing CS and SS, there was a three-way length X delay X participant interaction in repetition because RTs increase for SS at



A: Effects of Frequency, Imageability and Delay on Onset RTs.



**Figure 3.** Onset RT in computerized tasks: Effects of frequency, imageability, length and delay. A: Effects of Frequency and Imageability and Delay on word durations. B: Effects of Length and Delay on word durations.

delay > 0 (not expected if plans can be stored) but did not systematically change anywhere for CS.

**Conclusion.** As expected, all participants and all tasks showed significant benefits of preparation. Also, as expected, in the younger controls, onset RTs were longer for low frequency/low imageability words and any disadvantage was attenuated with delay. In CS, there was only a frequency effect when there was no delay in repetition, but this is in the context of significant frequency effects in reading and naming errors. Finally, across participants there was no consistent pattern indicating a length effect present at 0 delay that disappears with the ability to prepare.

Results show that our materials are sensitive for detecting effects of psycholinguistic variables. Importantly, they show that preparation is beneficial for stages of speech processing preceding articulation. The possibility to prepare benefits lexical access, as shown by a reduction of frequency and imageability effects with delay. However, if the

benefits of preparation extend to articulatory planning, they should also extend to analyses of word durations. If articulatory plans can be stored, there should be no effects of frequency on word duration after a delay since a plan will have been compiled and stored beforehand. Instead, if articulatory planning occurs online (during production), low-frequency words may show longer durations since their planning will require more time and this should not reduce with preparation. Effects of length are a given because it takes longer to say a longer word. However, if planning occurs mostly offline, these effects should reduce with preparation and this did not happen.

### Word durations

Figure 4 shows results for word durations with Panel A showing effects of word frequency and imageability, and Panel B showing effects of length. Corresponding statistical analyses are reported in Tables 5 and 6.

## B: Effects of Length and Delay on Onset RT

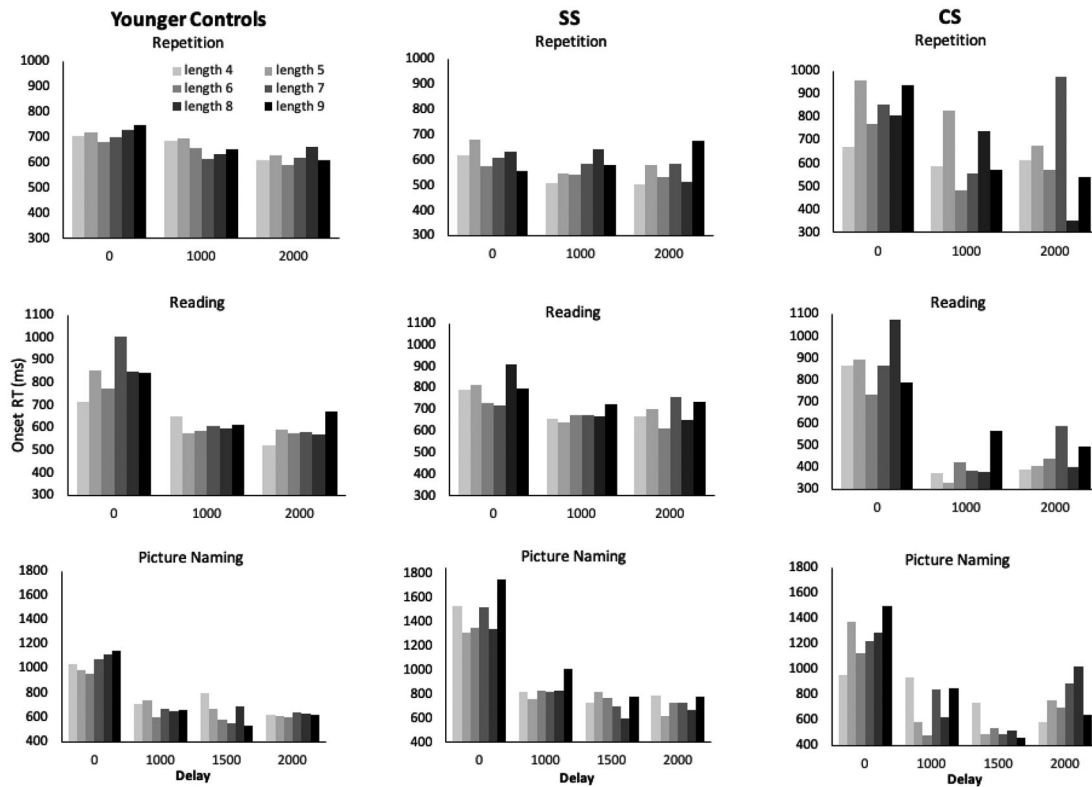


Figure 3 Continued

**Effect of participant (CS vs SS)**

CS produced words more slowly than SS. This was significant with the frequency lists in repetition ( $G^2(1) = 31.7$ ,  $p < .001$ , marginal  $R^2 = 5.3$ ), reading ( $G^2(1) = 166$ ,  $p < .001$ , marginal  $R^2 = 24.6$ ) and naming ( $G^2(1) = 82.4$ ,  $p < .001$ , marginal  $R^2 = 38.9$ ) and with the length list in reading ( $G^2(1) = 27.3$ ,  $p < .001$ , marginal  $R^2 = 11.6$ ) and naming ( $G^2(1) = 115$ ,  $p < .001$ , marginal  $R^2 = 27.1$ ), failing to reach significance only in repetition. Note, however, that in repetition CS made many errors and repeated attempts, which are not considered here. These results are consistent with CS having a deficit in articulatory planning.

**Effects of frequency, imageability and length**

Both younger controls and CS produced high-frequency words more quickly than low-frequency words in all tasks (effects are marginal in CS in naming) but the effect was significant only after a delay (see General Discussion). There were also effects of imageability in the younger controls, in reading and repetition with and without a delay. SS

showed only a marginal effect of frequency after a delay in reading, but in a single participant with high proficiency effects will be more difficult to detect. Overall, results were not consistent with articulatory plans that can be prepared and then stored for production. If a plan could be compiled before articulation starts, variables reflecting practice in articulation (frequency) and/or phonological encoding (imageability) should not affect durations or their effect should reduce with preparation.

Strong and significant effects of length on word duration were shown in all tasks and in all participants. This is expected since longer words will take longer to say.

**Effects of delay**

There were no positive effects of delay on word durations across participants, type of list (frequency and length) and task (repetition, reading and naming). The opportunity to prepare did not increase the speed at which words were produced. Importantly, this was the case both for control participants and for CS, whose word durations are significantly

**Table 3.** Analyses of onset RT for lists contrasting frequency, imageability and delay.

			Younger controls		SS		CS		CS vs SS	
			effect size	p-value	effect size	p-value	effect size	p-value	effect size	p-value
<b>Frequency</b>	<b>Delay = 0</b>	Repetition	<b>0.3</b>	.05	–	–	–	–	–	–
		Reading	<b>0.4</b>	.04	–	–	–	–	–	–
		Naming	–	–	–	–	–	–	–	–
	<b>Delay &gt; 0</b>	Repetition	–	–	–	–	<b>5.1</b>	.05	<b>2.7</b>	.001
		Reading	–	–	–	–	–	–	<b>1</b>	.03
								Larger freq effect in CS		
								Larger freq effect in CS		
			<b>0.7</b>	.005	–	–	–	–	–	–
			LF words faster							
<b>Imageability</b>	<b>Delay = 0</b>	Repetition	–	–	–	–	<b>11.5</b>	.03	–	–
		Reading	<b>0.04</b>	.03	–	–	–	–	–	–
	<b>Delay &gt; 0</b>	Repetition	–	–	–	–	–	–	–	–
<b>Frequency X Imageability</b>	<b>Delay = 0</b>	Repetition	<b>0.2</b>	.08	–	–	–	–	–	–
		Reading	<b>0.6</b>	.007	–	–	–	–	–	–
			Low freq/low imag words slower							
<b>Delay</b>	<b>0 vs &gt;0</b>	Repetition	–	–	–	–	–	–	–	–
		Reading	–	–	–	–	–	–	–	–
		Naming	–	–	–	–	–	–	–	–
								Participant X Delay		
			<b>2.7</b>	<.001	<b>8.9</b>	<.001	<b>30.2</b>	<.001	<b>3.8</b>	<.001
								More benefit of delay in CS		
			<b>10.5</b>	<.001	<b>6.2</b>	.003	<b>28.4</b>	<.001	<b>6.9</b>	<.001
								More benefit of delay in CS		
			<b>14.2</b>	<.001	<b>42.4</b>	<.001	<b>26</b>	<.001	–	–

Only statistically significant effect sizes are reported based on marginal R squared values from a generalized linear model. Freq. = frequency; imag. = imageability. Effects of frequency and imageability in the expected direction: slower RTs for low frequency/low imageability words.

slower. In fact, there were some significant effects in the opposite direction. In the younger controls, durations were longer with delay on the frequency list in reading, and on the length lists in all tasks. In CS, there was a marginal paradoxical effect of delay in repetition on the frequency list (durations longer after a delay), which was stronger than in SS. We will address these results in the General Discussion.

#### Interactions: length X delay

Interactions between length and delay were inconsistent. In the younger controls, delay reduced the length effect in reading, but because the shorter words slowed down with delay, not because the longer words became faster. In CS, a significant length X delay interaction in naming was created only by long durations with no delay only when length = 8. In reading and repetition there were no interactions and there were no interactions in SS. The interaction of length X delay X participant when comparing CS to SS in naming was due to the interaction we have already noted in CS, which was not apparent in SS.

## Discussion

We asked a group of young neurotypical speakers, a speaker with a form of AoS, and an age-matched control to produce words (repeat, read or name pictures) immediately or after a go signal (delayed conditions). We measured both onset RT and word durations. We assumed that onset RTs reflect lexical access and phonological encoding with a possible contribution of articulatory planning. We assumed, instead, that word durations tap subsequent processes, especially online articulatory planning. We also assumed that if lexical, phonological and articulatory processes can be prepared during a delay, onset RTs will get faster (see also Balota & Chumbley, 1985; Buz & Jaeger, 2016). Instead, effects on word durations will depend on how articulatory planning is carried out.

We contrasted two hypotheses. One possibility is that articulatory planning is carried out offline: a programme corresponding to at least a word must be prepared and made ready in an articulatory buffer before articulation can start (Meyer & Schriefers, 1991; Pierrehumbert, 2002; Roelofs, 2002a). A second possibility is that articulatory planning is mostly carried out online



## B: Effects of Length and Delay on word durations

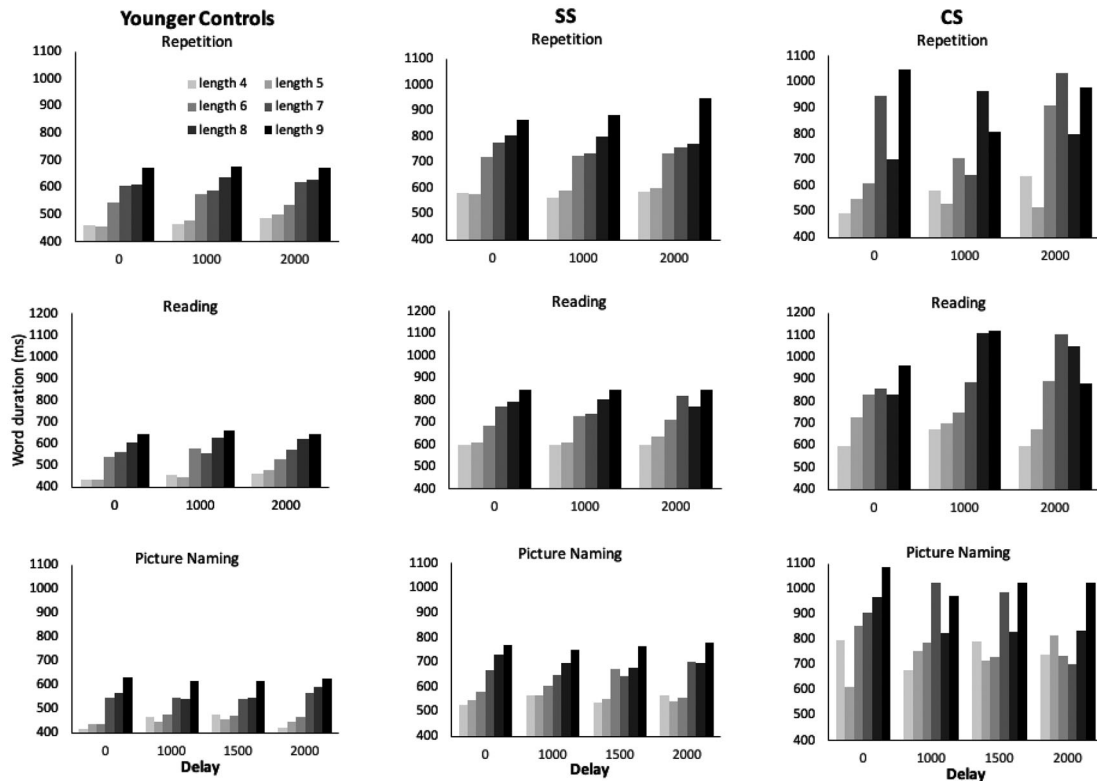


Figure 4. Continued

Both hypotheses predict an effect of preparation on onset RTs. However, only the hypothesis that articulatory planning occurs offline predicts a reduction in word durations in delayed conditions when an articulatory plan can be prepared. In contrast, only the hypothesis that articulatory planning occurs online predicts significant effects of frequency and, possibly, imageability while articulation is carried out, that is, on word durations. Moreover, these effects, like length effects, should not reduce with preparation. Only according to this second hypothesis, in fact, does one expect psycholinguistic variables to affect speech after the beginning of articulation. Table 7 summarizes our results in relation to our predictions.

With onset RTs we found:

- Strong effects of preparation, with faster RTs in delayed conditions across lists, tasks and participants.
- In control participants, significant effects of frequency and imageability which weakened with delay (although not in all tasks).
- In control participants, a significant length effect only in naming which became paradoxical with

delay (faster RTs for longer words). Other results were marginal and/or inconsistent

These results confirm the sensitivity of our tasks and conditions to our experimental manipulations. The lack of length effects in onset RTs is more consistent with articulatory planning occurring online since preparation of longer words should have delayed speech should have taken longer. Word durations are more selectively associated with articulatory planning and provide additional crucial data to distinguish the hypotheses in hand. With word durations we found:

- No positive effect of preparation across lists, tasks and participants. In fact, there were significant effects in the opposite direction, with durations being longer after delay in several conditions. No positive effect of preparation on CS's dysfluencies and errors, including repeated attempts.
- A significant effect of frequency in control participants and CS across tasks, with a larger frequency effect in CS than in his matched control.

**Table 5.** Analyses of word durations for lists contrasting frequency, imageability and delay.

			Younger Controls		SS		CS		CS vs SS	
			effect size	p-value	effect size	p-value	effect size	p-value	effect size	p-value
<b>Frequency</b>	<b>Delay = 0</b>	Repetition	–	–	–	–	–	–	–	–
		Reading	–	–	–	–	–	–	–	–
		Naming	–	–	–	–	–	–	–	–
	<b>Delay &gt; 0</b>	Repetition	<b>2.2</b>	<.001	–	–	<b>7.8</b>	.01	<b>2.5</b>	<.001
								Larger freq effect in CS		
		Reading	<b>2.2</b>	<.001	<b>3.3</b>	.07	<b>14.5</b>	.003	<b>2.9</b>	<.001
								Larger freq effect in CS		
<b>Imageability</b>	<b>Delay = 0</b>	Naming	<b>0.5</b>	.006	–	–	<b>7.6</b>	.07	–	–
		Repetition	<b>1.4</b>	<.001	–	–	–	–	–	–
		Reading	<b>1.1</b>	<.001	–	–	–	–	–	–
	<b>Delay &gt; 0</b>	Repetition	<b>0.3</b>	.01	–	–	–	–	–	–
		Reading	<b>0.2</b>	.04	–	–	–	–	–	–
<b>Frequency X Imageability</b>	<b>Delay = 0</b>	Repetition	–	–	–	–	–	–	–	–
		Reading	–	–	–	–	–	–	–	–
	<b>Delay &gt; 0</b>	Repetition	–	–	–	–	–	–	–	–
		Reading	–	–	–	–	–	–	–	–
<b>Delay</b>	<b>0 vs &gt;0</b>	Repetition	–	–	–	–	<b>2.7</b>	.06	<b>1.1</b>	.008
							longer duration after delay			
		Reading	<b>0.3</b>	.001	–	–	–	–	–	–
								longer duration after delay		
		Naming	–	–	–	–	–	–	–	–

Only statistically significant effect sizes are reported based on marginal R squared values from a generalized linear model. Marginally significant results are in grey.

**Table 6.** Analyses of word durations for lists contrasting length and delay.

			Younger Controls		SS		CS		CS vs SS	
			effect size	p-value	effect size	p-value	effect size	p-value	effect size	p-value
<b>Length</b>	<b>Delay = 0</b>	Repetition	<b>35.4</b>	<.001	<b>74.4</b>	<.001	<b>34.1</b>	.02	–	–
		Reading	<b>33.9</b>	<.001	<b>66.6</b>	<.001	<b>43.9</b>	.002	<b>2.1</b>	.02
								stronger length in CS		
	<b>Delay &gt; 0</b>	Naming	<b>30.1</b>	<.001	<b>64.4</b>	<.001	<b>41.9</b>	<.001	–	–
Repetition		<b>30.8</b>	<.001	<b>66.4</b>	<.001	<b>30.1</b>	<.001	–	–	
Reading		<b>29.4</b>	<.001	<b>58.5</b>	<.001	<b>37.5</b>	<.001	<b>3.1</b>	<.001	
							slightly flatter for delay>0 (shorter words slower)			
<b>Delay</b>	<b>0 vs &gt;0</b>	Naming	<b>20.7</b>	<.001	<b>43.4</b>	<.001	<b>18.2</b>	<.001	–	–
		Repetition	<b>0.3</b>	.008	–	–	–	–	–	–
		Reading	<b>30.7</b>	<.001	–	–	–	–	–	–
		Naming	<b>0.4</b>	.004	–	–	–	–	–	–
							for all tasks durations longer after a delay			

Only statistically significant effect sizes are reported based on marginal R squared values from a generalized linear model. Marginally significant results are in grey.

- (c) In controls, a significant imageability effect, although smaller than the frequency effect.
- (d) No disappearance of the frequency effect in delayed conditions. In fact, the opposite was true, with significant effects of frequency only in delayed conditions both in controls and in CS.
- (e) Only marginal and inconsistent weakening of length effects for delays > 0 (for consistent results see Damian et al., 2010; for reading see Marinelli et al., 2016; Spinelli et al., 2005).

These findings are in strong contrast with articulatory plans being completed before speech initiation, not only because of lack of preparation effects but because of significant effects of frequency and imageability. This could be due to high-frequency/high-imageability words being associated with more practised plans. It is also possible that high-imageability words are able to send a stronger phonological signal from the buffer to the planner and this reduces any residual phonological encoding which is carried out online. In



**Table 7.** Summary of predictions of different hypothesis. Darker shading and ticks indicate fulfilled predictions; x indicates unfulfilled prediction. Lighter shading is for predictions that do not contrast.

	ARTICULATION OFFLINE & PREPARED IN A BUFFER					ARTICULATION ONLINE & NOT PREPARED LIKE PHONOLOGY				
	Onset RTs		Word durations		Errors	Onset RTs		Word durations		Errors
	Controls	CS	Controls	CS	CS	Controls	CS	Controls	CS	CS
<b>Delayed better than immediate?</b>										
Repetition	Yes: faster✓	Yes: faster✓	yes: shorter x	shorter x	fewer x	Yes: faster✓	Yes: faster✓	No: ✓same	No: ✓(longer)	no change ✓
Reading	Yes: faster✓	Yes: faster✓	yes: shorter x	shorter x	fewer x	Yes: faster✓	Yes: faster✓	No: ✓(longer)	No: ✓same	no change ✓
Naming	Yes: faster✓	Yes: faster✓	yes: shorter x	shorter x	fewer x	Yes: faster✓	Yes: faster✓	No: ✓same	No: ✓same	no change ✓
<b>High frequ/imm words better than low?</b>										
Repetition	Yes: faster✓	faster✓	No same x	No same x	X	Yes: faster✓	faster✓	Yes: shorter✓	Yes: shorter✓	X
Reading	Yes: faster✓	faster✓	No same x	No same x	fewer ✓	Yes: faster✓	faster✓	Yes: shorter✓	Yes: shorter✓	Fewer ✓
Naming	Yes: faster x*	Yes: faster x	No same x	No same ✓	X	Yes: faster x *	Yes: faster x	Yes: shorter✓	Yes: shorter x	X
* but fewer errors										
<b>Any frequ/imm effect, weaker in delayed conditions?</b>										
Repetition	Yes: weaker✓	Yes: weaker✓	yes weaker x	yes weaker x		Yes weaker ✓	Yes weaker ✓	No: ✓(stronger)	No: ✓(stronger)	
Reading	Yes: weaker✓	Yes weaker x	yes weaker x	yes weaker x		Yes weaker ✓	Yes weaker x	No: ✓(stronger)	No: ✓(stronger)	
Naming	Yes: weaker✓	Yes weaker x	yes weaker x	yes weaker x		Yes weaker ✓	Yes weaker x	No: ✓(stronger)	No: ✓(stronger)	
<b>Longer worse than shorter words</b>										
Repetition	Yes slower x	Yes slower x	Yes: longer ✓	Yes: longer ✓		No same ✓	No same x	Yes: longer ✓	Yes: longer ✓	
Reading	Yes slower	Yes slower x	Yes: longer ✓	Yes: longer ✓		No same ✓	No same ✓	Yes: longer ✓	Yes: longer ✓	
Naming	Yes: slower ✓	Yes slower x	Yes: longer ✓	Yes: longer ✓		No same x	No same ✓	Yes: longer ✓	Yes: longer ✓	
<b>Any length effect, weaker in delayed conditions?</b>										
Repetition	Yes weaker: x no or paradoxical		Yes: weaker x	Yes: weaker x		No ✓ no or paradoxical		No: similar✓	No: similar✓	
Reading			Yes: = marginal	Yes: weaker x	No: similar✓			No: similar✓		
Naming			Yes: = marginal	Yes: weaker ✓	No: similar✓			No: weaker x		

Text show predictions; shading and tick indicate that prediction is fulfilled; x indicates prediction not fulfilled. Lighter shading is for predictions that do not contrast.

any case, a frequency/imageability effect on word durations strengthens the view that a complete articulatory plan is not prepared and buffered before production. Moreover, some of these effects, rather than weakening with preparation, were stronger in delayed conditions (see later for more discussion).

In our previous investigation we argued that CS suffered from an impairment in the transfer of phonological information to the articulatory planner and for this reason his speech was marred by dysfluencies (syllabifications), phonetic and phonological errors, and especially repeated attempts, with false starts often followed by correct and fluent production. The results presented here confirm our interpretation. His word durations were significantly longer than those of an age- and education-matched control speaker and he made a high proportion of repeated attempts across all three production tasks. A reduction in a phonological output buffer has been suggested as an explanation of some features of apraxia of speech, such as syllabifications and reduced coarticulation (e.g., Rogers & Storkel, 1999). This hypothesis has been criticized because it predicts phonological but not phonetic errors, which are a hallmark of AoS (see also Miller & Guenther, 2021). However, if we hypothesize difficulties transferring information from the buffer to the articulatory planner, we predict phonological errors, phonetic errors and dysfluencies, as was found in CS and in the patients described by Rogers and Storkel (1999). All these errors will arise because information is not fed quickly and smoothly enough to the articulatory planner. This locus of impairment would define a particular variety of AoS, to be distinguished from other varieties where there is disruption to the spatio-temporal parameters needed to realize articulatory gestures or difficulties in their implementation (e.g., see Galluzzi et al., 2015).

Given his impairment, it is very significant that CS showed no advantage of preparation on word durations, on overall number of errors and, more specifically, on repeated attempts. One could argue that CS did not show any preparation advantages because his impairment reduced the capacity of an articulatory buffer to such an extent that no preparation was possible. However, since control speakers also showed no benefits of preparation, and, if anything, showed effects in the opposite direction, this is unlikely. Thus, our case is strengthened by converging results

that come from control speakers and from an aphasic speaker.

In designing our experimental investigation, we worried that some of the predictions from the hypothesis of online articulatory planning relied on null results. Word production is a very well-practised task and differences in word durations could be difficult to detect in proficient speakers. For this reason, it was important to have results from both control participants and an aphasic participant with a deficit that affected articulatory programming. Our results, however, rule out a lack of sensitivity in our measures. With word durations, no benefit of preparation contrasts with significant advantages for high frequency and high imageability words, in control participants and/or in CS. Crucially, not only did we fail to find benefits of preparation, we found small but highly significant effects in the opposite direction. This occurs with the frequency lists in reading, and with the length lists in all tasks, demonstrating there was no problem of sensitivity. Word durations are clearly sensitive to different effects and can shed light on the processing carried out after the beginning of articulation.

Negative effects of preparation on word durations were not expected, but they also showed that phonological and articulatory processes are not completed before articulation starts. Instead, it is possible that processing can be flexibly allocated offline or online depending on task requirements. In our task, a “go” signal, which appears after a delay, may put extra pressure on participants to articulate right away what has been prepared. This, however, may mean that some of the processes which would normally be carried out before the start of articulation are moved online (for consistent results see Damian, 2003; Kawamoto et al., 2014). This might also explain why frequency affects word durations more strongly after a delay. As more processes are moved online, any difference in the ease of articulatory planning will become more pronounced. This explanation is consistent with evidence from other paradigms where effects of psycholinguistic variables on durations become more evident when speakers are under pressure (see Kello et al., 2000, 2004), but note that these effects are inconsistent across studies (see Fink et al., 2018; Goldrick et al., 2019).

Trade-offs between onset RTs and word durations have been found in other studies. Holbrook et al. (2019), in a naming task, asked half of their

participants to begin speech as soon as possible and the other half to keep speech as brief as possible. They found that, when primed with the first phoneme of the word, participants in the first group started speech sooner, but increased word durations in a compensatory fashion. It is possible that the processes which are shifted online during word production involve phonological encoding. For example, Schriefers and Teruel (1999) showed that in “hasty” speakers only the first syllable of a word primed onset RTs, indicating that only this syllable was phonologically encoded before production. In contrast, non-hasty speakers demonstrated priming for both syllables. It is possible that the go signal influences the size of the articulatory planning unit that is used for articulation. When more time is available, a larger articulatory unit can be selected. When the planner must respond quickly, a smaller articulatory unit is selected so that articulation can start sooner, but then more processing must be carried out online during speech.

The possibility that speech gestures of different sizes can be selected according to demands is not new in production models. In the Selection-Coordination model (Tilsen, 2013, 2016), phonological units are activated in parallel but serially selected according to an activation gradient. As language competence grows, more units are co-selected to activate coordinated gestures which are temporally coupled with oscillators for execution. These coordinated sets of gestures are generally syllable-sized in adult speakers, but the ability to use smaller units according to task demands would be maintained. Similarly, Krause and Kawamoto (2020) hypothesize a context-sensitive gating process. When this process operates with smaller units, there will be less coarticulation and longer word durations. Our results are consistent with these hypotheses.

Our results are only in apparent contradiction with other results from the literature. Some study-time preparation effects, previously attributed to articulatory planning, can easily be attributed to earlier levels since they are based on onset RTs (as discussed in the Introduction; e.g., see Maas & Mailend, 2012). Results showing priming of different kinds of syllabic units are equally susceptible to alternative interpretations. Some studies have found that syllable primes facilitate word onset RTs even when they correspond to a non-initial syllable of the target (Meyer &

Schriefers, 1991; Roelofs, 2002a). However, these effects do not occur in all speakers, as mentioned above (see Schriefers & Teruel, 1999), are not always replicated (see Roelofs, 2004), and, most importantly, can be interpreted as phonological and attentional effects. For example, Roelofs (2002a) showed that a second-syllable prime speeds up production even with a homogeneous set of response words, where all the words start with the same syllable. According to Roelofs, this showed that planning for the whole word must be completed before speech starts. Otherwise, onset RTs would not be affected when the first syllable of the word is known in advance and planning of the second syllable can occur online. However, the second-syllable primes may have facilitated phonological encoding and preparation in a phonological buffer rather than articulatory planning. Moreover, the primes did not show facilitation, but only reduced interference compared to a neutral condition, suggesting that the main impact of the primes was to modulate the attention of the participants, not to change phonological/articulatory planning. Word durations offer a better measure of articulatory planning, less contaminated by phonological influences.

Our results are also in only apparent contrast with those of Kello and Plaut (2000) and Kello (2004) who found faster onset RTs and shorter durations with more pressure to respond. Kello et al. used a “tempo-naming” task where participants were exposed to the beats of a metronome and then had to name a (printed) word in synchronization with the last beat. A visual display helped setting up the tempo and indicating how fast it would be. In these conditions, participants were encouraged to adjust their processing rate according to the tempo, explaining why quicker tempos decreased both onset RTs and word durations and why this manipulation was effective even in “delayed” conditions where the printed word was on view from the beginning of each trial. In our experiments, instead, participants had no incentive or possibility to adjust their processing rate because the “go” signal appeared at unpredictable delays after stimulus presentation. There was no set rhythm to modulate speed of articulation. Therefore, in delayed conditions participants had more time to prepare the phonology of the word and when the “go” signal appeared they felt more pressure to start articulation. However, a quick start of articulation could be accompanied by longer

durations because more processing is shifted online. With a paradigm similar to ours, Munson (2007) also found longer vowel durations in delayed than immediate naming.

Taken together, our results provide strong evidence that articulatory planning (and possibly some phonological processing) occurs online and that the final stage of speech production—articulatory/motor planning/programming—cannot be prepared and buffered in the same way as previous stages can. Buffered phonological representations are versatile and available to consciousness. They can be refreshed and they are used in a variety of tasks such as repeating lists of words, writing to dictation or decoding a sentence heard in a noisy environment (Baddeley & Hitch, 2019). They can be manipulated in spoonerisms and phoneme deletion tasks. Consistent with other studies, we have demonstrated that a preparation delay between presentation of a stimulus and a go signal greatly reduces onset RTs (e.g., Balota & Chumbley, 1985; Kawamoto et al., 2008; Laganaro & Alario, 2006; Mooshammer et al., 2009, 2012). Articulatory gestures do not demonstrate the same characteristics. They appear inaccessible to our introspection and our experimental investigation shows that they cannot be prepared in the same way.

Articulatory plans of various sizes may be pre-packaged and stored for use by the articulatory system, but, once retrieved, they are executed in strict sequential fashion, with any memory resources encapsulated within the system. Allowing a delay in responding does not help. The articulatory planner is only engaged during actual articulation, which must start and run from beginning to end. To use a sports metaphor, the amount of time a skier waits at the starting line will not help her to complete a race better or faster. This does not mean that, once started, the process of articulatory production cannot be modulated by incoming feedback, or that it is not subject to conscious control. In fact, we have presented some evidence for flexibility in the process, with trade-offs between onset RTs and word durations (for discussion and consistent evidence see Krause & Kawamoto, 2020; see also Schmidt, 1975; Schmidt et al., 2019, on the need for adjustments depending on phonological and communicative context). What we argue, however, is that we cannot practise articulatory processes without engaging them, because motor

representations are not available for mental manipulation in the same way that phonological or visual representations are. Thus, mental preparation will not help, only motor practice will. To continue with our metaphor, having done the slope before will help a skier to complete it faster a second time. Similarly, practice will help with articulatory planning and we have shown that high-frequency words have shorter durations than low frequency words in control speakers and elicit fewer errors in CS.

### Conclusions

Our results show a strong contrast between effects in onset RTs, tapping processes preceding the start of articulation, and word durations, tapping online processes during articulation. From a theoretical point of view, our results show that articulatory representations cannot be buffered and prepared as phonological representations can (there are no effects of preparation on word duration). Articulatory plans do not need to be fully compiled or buffered before articulation starts. Instead, articulatory processing and, possibly, some phonological processing occurs online during word production, as demonstrated by frequency and imageability effects on word durations. Models of speech production like the DIVA/GODIVA model (Miller & Guenther, 2021) are not currently able to represent a contrast between a phonological buffer and encapsulated resources supporting articulatory planning.

From a methodological point of view, our results demonstrate the importance of using word durations as a sensitive measure of articulatory fluency and also illustrate the danger of assuming that articulatory planning is completed before articulation starts when naming is measured with onset RTs in delayed conditions (e.g., see Laganaro & Alario, 2006). They also demonstrate the advantage of carrying out chronometric analyses of speech that include both neurotypical and aphasic participants. In our study, combined analyses helped to strengthen our theoretical claims while also confirming previous conclusions about the articulatory nature of CS's speech impairment.

Finally, our results suggest that rehabilitation of articulatory deficits may only succeed when it actually engages spoken production. One cannot mentally practise articulation, for example, by reading, if not reading aloud.

## Notes

1. Note that all models make exactly the same assumptions regarding the different stages involved in speech production after lexical access (stage 2 above). In some models, the purpose of phonological encoding is to link phonemes to syllable structure (e.g., Levelt et al., 1999; Roelofs, 1996, 1997), while, in other models, phonemes are already organized into syllables at the lexical level since a syllabic organization is crucial for optimal storage and retrieval of information (e.g., see Romani et al., 2011). In models where phonemes are associated with syllables post-lexically, there is no clearly declared phonological output buffer, but filled syllabic frameworks might serve this function (see also Dell et al., 1993). In the original model by Levelt et al. (1999), an (articulatory) output buffer was placed *after* phonological encoding, at the level where articulatory syllable plans/gestures are retrieved and sequenced for production.
2. But there is some question about whether these effects are due to imageability, age of acquisition, or frequency given the overlap between these variables (see Cortese et al., 2018; Ellis & Monaghan, 2002).
3. Note that it is logically possible for an articulatory buffer to be used only optionally. However, the delayed naming conditions of our investigation are ideal to prepare and store articulatory plans. If an articulatory buffer is not used in these conditions, it is difficult to envision conditions in which it would be.
4. Different cells need to be in different colours for the table to be understandable.

## Acknowledgments

This work was supported by a Ph.D. studentship to Dinesh Ramoo granted by the University of Birmingham and by a master's dissertation carried out by Priya Silverstein, also at the University of Birmingham. We would like to thank Dawn Jevons from the Stroke Association in South Birmingham for her assistance in recruiting the aphasic participant. We are extremely grateful to CS and his wife for many hours of graceful engagement. We are also grateful to the EPS small grant scheme for the support we received to complete the study (grant to Cristina Romani).

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by a Ph.D. studentship to Dinesh Ramoo granted by the University of Birmingham and by a master's dissertation carried out by Priya Silverstein, at Aston

University ; and by a small grant from the Experimental Psychology Society.

## ORCID

Cristina Romani  <http://orcid.org/0000-0002-5693-4131>

## References

- Agresti, A. (2013). *Categorical data analysis* (3rd ed). Wiley.
- Arnold, J E. (2016). Explicit and emergent mechanisms of information status. *Topics in Cognitive Science*, 8, 737–760.
- Baayen, R., Piepenbrock, R., & Rijn, H. (1993). *The CELEX lexical data base*.
- Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 39 (3), 331–346. <https://doi.org/10.1006/jmla.1998.2572>
- Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, 112, 91–106. <https://doi.org/10.1016/j.cortex.2018.05.015>
- Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, 24 (1), 89–106. [https://doi.org/10.1016/0749-596X\(85\)90017-8](https://doi.org/10.1016/0749-596X(85)90017-8)
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133 (2), 283–316. <https://doi.org/10.1037/0096-3445.133.2.283>
- Barton, J. J. S., Hanif, H. M., Björnström, L. E., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive Neuropsychology*, 31(5–6), 378–412. <https://doi.org/10.1080/02643294.2014.895314>
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9), 341–345.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3–4), 155–180. <https://doi.org/10.1159/000261913>
- Buxo-Lugo, A., Jacobs, C. L., & Watson, D. G. (2020). The world is not enough to explain lengthening of phonological competitors. *Journal of Memory and Language*, 110, 104066. <https://doi.org/10.1016/j.jml.2019.104066>
- Buz, E., & Jaeger, T. F. (2016). The (in)dependence of articulation and lexical planning during isolated word production. *Language, Cognition and Neuroscience*, 31(3), 404–424. <https://doi.org/10.1080/23273798.2015.1105984>
- Code, C. (1998). Models, theories and heuristics in apraxia of speech. *Clinical Linguistics & Phonetics*, 12(1), 47–65. <https://doi.org/10.3109/02699209808985212>
- Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, 33 (4), 497–505. <https://doi.org/10.1080/14640748108400805>
- Cortese, M. J., Yates, M., Schock, J., & Vilks, L. (2018). Examining word processing via a megastudy of conditional reading



- aloud. *Quarterly Journal of Experimental Psychology*, 71(11), 2295–2313. <https://doi.org/10.1177/1747021817741269>
- Croot, K., Lalas, G., Biedermann, B., Rastle, K., Jones, K., & Cholin, J. (2017). Syllable frequency effects in immediate but not delayed syllable naming in English. *Language, Cognition and Neuroscience*, 32(9), 1119–1132. <https://doi.org/10.1080/23273798.2017.1284340>
- Damian, M. F. (2003). Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(3), 416–431. <https://doi.org/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. <https://doi.org/10.1037/a0019446>
- Deger, K., & Ziegler, W. (2002). Speech motor programming in apraxia of speech. *Journal of Phonetics*, 30(3), 321–335. <https://doi.org/10.1006/jpho.2001.0163>
- de Groot, A. (n.d.). Representational aspects of word imageability and word frequency as assessed through word association. <https://oae.ovid.com/article/00004786-198909000-00006/HTML>
- 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(2), 149–195. [https://doi.org/10.1207/s15516709cog1702\\_1](https://doi.org/10.1207/s15516709cog1702_1)
- Ellis, A. W., & Monaghan, J. (2002). Reply to Strain, Patterson, and Seidenberg (2002). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1), 215–220. <https://doi.org/10.1037/0278-7393.28.1.215>
- Fink, A., Oppenheim, G. M., & Goldrick, M. (2018). Interactions between lexical access and articulation. *Language, Cognition and Neuroscience*, 33(1), 12–24. <https://doi.org/10.1080/23273798.2017.1348529>
- Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3), 474–496. <https://doi.org/10.1353/lan.0.0035>
- Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4), 789–806. <https://doi.org/10.1016/j.jml.2011.11.006>
- Galluzzi, C., Bureca, I., Guariglia, C., & Romani, C. (2015). Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. *Neuropsychologia*, 71, 64–83. <https://doi.org/10.1016/j.neuropsychologia.2015.03.007>
- Goldrick, M., McClain, R., Cibelli, E., Adi, Y., Gustafson, E., Moers, C., & Keshet, J. (2019). The influence of lexical selection disruptions on articulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(6), 1107–1141. <https://doi.org/10.1037/xlm0000633>
- Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition*, 102(2), 219–260. <https://doi.org/10.1016/j.cognition.2005.12.010>
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13(2), 135–145. <https://doi.org/10.1038/nrn3158>
- Holbrook, B. B., Kawamoto, A. H., & Liu, Q. (2019). Task demands and segment priming effects in the naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(5), 807–821. <https://doi.org/10.1037/xlm0000631>
- Jacobs, C. L., Yiu, L. K., Watson, D. G., & Dell, G. S. (2015). Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. *Journal of Memory and Language*, 84, 37–48. <https://doi.org/10.1016/j.jml.2015.05.004>
- Jordan, M. I. (1990). Motor learning and the degrees of freedom problem. (By M. Jeannerod; pp. 796–836). LEA.
- Kahn, J. M., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience*, 30(1–2), 103–119. <https://doi.org/10.1080/01690965.2013.848989>
- Kawamoto, A. H., Kello, C. T., Higareda, I., & Vu, J. V. Q. (1999). Parallel processing and initial phoneme criterion in naming words: Evidence from frequency effects on onset and rime duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(2), 362–381. <https://doi.org/10.1037/0278-7393.25.2.362>
- Kawamoto, A. H., Liu, Q., Lee, R. J., & Grebe, P. R. (2014). The segment as the minimal planning unit in speech production: Evidence based on absolute response latencies. *Quarterly Journal of Experimental Psychology*, 67(12), 2340–2359. <https://doi.org/10.1080/17470218.2014.927892>
- Kawamoto, A. H., Liu, Q., Mura, K., & Sanchez, A. (2008). Articulatory preparation in the delayed naming task. *Journal of Memory and Language*, 58(2), 347–365. <https://doi.org/10.1016/j.jml.2007.06.002>
- Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology: Human Perception and Performance*, 30(5), 942–955. <https://doi.org/10.1037/0096-1523.30.5.942>
- Kello, C. T., & Plaut, D. C. (2000). Strategic control in word reading: Evidence from speeded responding in the tempo-naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 719–750. <https://doi.org/10.1037/0278-7393.26.3.719>
- Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech production. *Journal of Experimental Psychology: General*, 129(3), 340–360. <https://doi.org/10.1037/0096-3445.129.3.340>
- Klapp, S. T. (1995). Motor response programming during simple choice reaction time: The role of practice. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 1015–1027. <https://doi.org/10.1037/0096-1523.21.5.1015>
- Klapp, S. T. (2003). Reaction time analysis of two types of motor preparation for speech articulation: Action as a sequence of chunks. *Journal of Motor Behavior*, 35(2), 135–150. <https://doi.org/10.1080/00222890309602129>



- Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading: Reconsidered. *Journal of Experimental Psychology*, 100(2), 368–374. <https://doi.org/10.1037/h0035471>
- Kohn, S. E. (1984). The nature of the phonological disorder in conduction aphasia. *Brain and Language*, 23(1), 97–115. [https://doi.org/10.1016/0093-934X\(84\)90009-9](https://doi.org/10.1016/0093-934X(84)90009-9)
- Kohn, S. E. (1989). The nature of the phonemic string deficit in conduction aphasia. *Aphasiology*, 3, 209–239.
- Krause, P. A., & Kawamoto, A. H. (2020). On the timing and coordination of articulatory movements: Historical perspectives and current theoretical challenges. *Language and Linguistics Compass*, 14(6), <https://doi.org/10.1111/lnc3.12373>
- Laganaro, M., & Alario, F.-X. (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language*, 55(2), 178–196. <https://doi.org/10.1016/j.jml.2006.05.001>
- Lam, T. Q., & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory & Cognition*, 38(8), 1137–1146. <https://doi.org/10.3758/MC.38.8.1137>
- Lam, T. Q., & Watson, D. G. (2014). Repetition reduction: Lexical repetition in the absence of referent repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 829–843. <https://doi.org/10.1037/a0035780>
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, 22(1), 1–38; discussion 38–75. <https://doi.org/10.1017/s0140525x99001776>
- Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50(1–3), 239–269. [https://doi.org/10.1016/0010-0277\(94\)90030-2](https://doi.org/10.1016/0010-0277(94)90030-2)
- Lohmann, A. (2018). Time and thyme are NOT homophones: A closer look at Gahl's work on the lemma-frequency effect, including a reanalysis. *Language*, 94(2), e180–e190. <https://doi.org/10.1353/lan.2018.0032>
- Maas, E., & Mailend, M.-L. (2012). Speech planning happens before speech execution: Online reaction time methods in the study of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, 55(5), 1523–1534. [https://doi.org/10.1044/1092-4388\(2012/11-0311\)](https://doi.org/10.1044/1092-4388(2012/11-0311))
- Maas, E., Robin, D. A., Wright, D. L., & Ballard, K. J. (2008). Motor programming in apraxia of speech. *Brain and Language*, 106(2), 107–118. <https://doi.org/10.1016/j.bandl.2008.03.004>
- MacKay, I. R. A. (1987). *Phonetics: The science of speech production* (2nd ed). Little, Brown.
- Marinelli, C. V., Romani, C., Burani, C., McGowan, V. A., & Zoccolotti, P. (2016). Costs and benefits of orthographic inconsistency in reading: Evidence from a cross-linguistic comparison. *PLoS ONE*, 11(6), e0157457. <https://doi.org/10.1371/journal.pone.0157457>
- Mason, M. (1978). From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. *Memory & Cognition*, 6(5), 568–581. <https://doi.org/10.3758/BF03198246>
- Meyer, A. S., Belke, E., Häcker, C., & Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language*, 57(2), 210–231. <https://doi.org/10.1016/j.jml.2006.10.005>
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48(1), 131–147. [https://doi.org/10.1016/S0749-596X\(02\)00509-0](https://doi.org/10.1016/S0749-596X(02)00509-0)
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(6), 1146–1160. <https://doi.org/10.1037/0278-7393.17.6.1146>
- Miller, H. E., & Guenther, F. H. (2021). Modelling speech motor programming and apraxia of speech in the DIVA/GODIVA neurocomputational framework. *Aphasiology*, 35(4), 424–441. <https://doi.org/10.1080/02687038.2020.1765307>
- Mooshammer, C., Goldstein, L., Nam, H., McClure, S., Saltzman, E., & Tiede, M. (2012). Bridging planning and execution: Temporal planning of syllables. *Journal of Phonetics*, 40(3), 374–389. <https://doi.org/10.1016/j.wocn.2012.02.002>
- Mooshammer, C. R., Goldstein, L., Tiede, M., Kulshreshtha, M., McClure, S., & Katsika, A. (2009). Planning time effects of phonological competition: Articulatory and acoustic data. *The Journal of the Acoustical Society of America*, 125(4), 2657–2657. <https://doi.org/10.1121/1.4784180>
- Munson, B. (2007). Lexical access lexical representation and vowel production. *Laboratory Phonology*, 9, 201–228.
- Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, 51(6), 2533–2545. <https://doi.org/10.3758/s13428-018-1100-1>
- Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7*. De Gruyter Mouton. <https://doi.org/10.1515/9783110197105.1.101>
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77(2), 97–132. [https://doi.org/10.1016/S0010-0277\(00\)00090-1](https://doi.org/10.1016/S0010-0277(00)00090-1)
- Ramoo, D., Olson, A., & Romani, C. (2021). Repeated attempts, phonetic errors, and syllabifications in a case study: Evidence of impaired transfer from phonology to articulatory planning. *Aphasiology*, 35(4), 485–517. <https://doi.org/10.1080/02687038.2021.1881349>
- Roelofs, A. (1996). Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, 35(6), 854–876. <https://doi.org/10.1006/jmla.1996.0044>
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64(3), 249–284. [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9)
- Roelofs, A. (2002a). Spoken language planning and the initiation of articulation. *Quarterly Journal of Experimental Psychology – Human Experimental Psychology*, 55(2), 465–483. <https://doi.org/10.1080/02724980143000488>
- Roelofs, A. (2002b). Syllable structure effects turn out to be word length effects: Comment on Santiago et al. (2000). *Language and Cognitive Processes*, 17(1), 1–13. <https://doi.org/10.1080/01690960042000139>

- Roelofs, A. (2004). Comprehension-based versus production-internal feedback in planning spoken words: A rejoinder to Rapp and Goldrick (2004). *Psychological Review*, 111(2), 579–580. <https://doi.org/10.1037/0033-295X.111.2.579>
- Rogers, M. A., & Storkel, H. L. (1999). Planning speech one syllable at a time: The reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology*, 13(9–11), 793–805. <https://doi.org/10.1080/026870399401885>
- Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology*, 22(7), 817–850. <https://doi.org/10.1080/02643290442000365>
- Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive Psychology*, 62(2), 151–192. <https://doi.org/10.1016/j.cogpsych.2010.08.001>
- Romani, C., Galuzzi, C., Guariglia, C., & Goslin, J. (2017). Comparing phoneme frequency, age of acquisition, and loss in aphasia: Implications for phonological universals. *Cognitive Neuropsychology*, 34(7–8), 449–471. <https://doi.org/10.1080/02643294.2017.1369942>
- Santiago, J., MacKay, D. G., & Palma, A. (2002). Length effects turn out to be syllable structure effects: Response to Roelofs (2002). *Language and Cognitive Processes*, 17(1), 15–29. <https://doi.org/10.1080/01690960042000148>
- 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), 1–44. <https://doi.org/10.1080/016909600386101>
- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4), 225–260. <https://doi.org/10.1037/h0076770>
- Schmidt, R. A., Lee, T. D., Winstein, C. J., Wulf, G., & Zelaznik, H. N. (2019). *Motor control and learning: A behavioral emphasis* (6th ed.). Human Kinetics.
- Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, 11(1), 17–50. <https://doi.org/10.1080/713752301>
- Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 24(1), 114. <https://doi.org/10.2307/748013>
- Shields, L. W., & Balota, D. A. (1991). Repetition and associative context effects in speech production. *Language and Speech*, 34(1), 47–55. <https://doi.org/10.1177/002383099103400103>
- Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology*, 27(2), 217–235. [https://doi.org/10.1207/s15326942dn2702\\_2](https://doi.org/10.1207/s15326942dn2702_2)
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). Academic Press.
- Tilsen, S. (2013). A dynamical model of hierarchical selection and coordination in speech planning. *Plos One*, 8(4), e62800. <https://doi.org/10.1371/journal.pone.0062800>
- Tilsen, S. (2016). Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics*, 55, 53–77. <https://doi.org/10.1016/j.wocn.2015.11.005>
- Van Der Merwe, A. (2009). A theoretical framework for the characterization of pathological speech sensorimotor control. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed, pp. 3–18). Thieme Medical Publishers.
- Van Der Merwe, A. (2021). New perspectives on speech motor planning and programming in the context of the four-level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. *Aphasiology*, 35(4), 397–423. <https://doi.org/10.1080/02687038.2020.1765306>
- Varley, R., & Whiteside, S. P. (2001). What is the underlying impairment in acquired apraxia of speech? *Aphasiology*, 15(1), 39–49. <https://doi.org/10.1080/02687040042000115>
- Varley, R. A., Whiteside, S. P., & Luff, H. (1999). Apraxia of speech as a disruption of word-level schemata: Some durational evidence. *Journal of Medical Speech and Language Pathology*, 7, 127–132.
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology Section A*, 50(2), 439–456. <https://doi.org/10.1080/713755710>
- Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, 37(3), 356–381. <https://doi.org/10.1006/jmla.1997.2517>
- Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, 60(4), 502–529. <https://doi.org/10.1016/j.jml.2009.02.001>