Original Paper

Folia Phoniatrica et Logopaedica

Folia Phoniatr Logop 2017;69:191–208
DOI: 10.1159/000485657

Uniqueness Point Effects during Speech Planning in Adults Who Do and Do Not Stutter

Geoffrey A. Coalson[a] Courtney T. Byrd[b] Amanda Kuylen[a]

[a]Department of Communication Sciences and Disorders, Louisiana State University, Baton Rouge, LA, USA;
[b]Department of Communication Sciences and Disorders, The University of Texas at Austin, Austin, TX, USA

Keywords
Stuttering · Speech production · Preverbal monitor · Phonological encoding · Adults

Abstract

Background/Aims: Previous studies employing a variety of tasks have demonstrated that adults who stutter (AWS) present with phonological encoding differences compared to adults who do not stutter (AWNS). The present study examined whether atypical preverbal monitoring also influenced AWS performance during one such paradigm – the silent phoneme monitoring task. Specifically, we investigated whether monitoring latencies for AWS were accelerated after the word's uniqueness point – the phoneme that isolates the word from all lexical competitors – as observed for AWNS when monitoring internal and external speech. ***Methods:*** Twenty adults (10 AWS, 10 AWNS) completed a silent phoneme monitoring task using stimuli which contained either (a) early uniqueness points (EUP), (b) late uniqueness points, or (c) no uniqueness point (NUP). Response latency when identifying word-final phonemes was measured. ***Results:*** AWNS exhibited the expected uniqueness point effect when monitoring internal speech; word-final phonemes were accessed more rapidly for words with EUP than NUP. In contrast, AWS did not differ in the phoneme monitoring speed. That is, AWS did not exhibit the expected uniqueness point effects. ***Conclusion:*** Findings suggest that inefficient or atypical preverbal monitoring may be present in AWS and support theories that implicate the internal speech monitor as an area of deficit.

Introduction

Theoretical accounts of stuttering identify several factors that may cause or contribute to moments of disfluent speech, including less coordinated motoric speech production (e.g., [1, 2]), inefficient speech-language planning (e.g., [3, 4]), or a dynamic interaction between these factors in the presence of internal or external demands (e.g., [5, 6]). In terms of speech planning difficulties, a considerable amount of data implicate weaknesses in phonological encoding early in life as a key area of compromise (e.g., [7–9], for an opposing viewpoint, see [10, 11]), particularly for children whose stuttering persists into adulthood (e.g., [12–15]). These data are further supported by studies which indicate that many adults who stutter (AWS) perform more poorly than adults who do

KARGER

E-Mail karger@karger.com
www.karger.com/fpl

Geoffrey A. Coalson
Department of Communication Sciences and Disorders, Louisiana State University
63 Hatcher Hall
Baton Rouge, LA 70803 (USA)
E-Mail gcoals1@lsu.edu

not stutter (AWNS) when completing experimental tasks that rely on efficient phonological processing (e.g., nonword repetition [16–18]; silent error monitoring [19]; phoneme elision [16, 17]; word jumble tasks [20]; silent rhyme judgment [21, 22]; silent phoneme monitoring [23–25]). Although differences in the phonological abilities of AWS are not unequivocal (cf. [26–28]), there are significant data to suggest that phonological processing difficulties do not disappear with age and may serve as a valuable prognostic factor during clinical assessment (e.g., [14, 15]).

The experimental tasks used in these adult studies, however, also require participants to silently monitor, evaluate, or manipulate internal speech. As suggested by Levelt et al. [29], both efficient *monitoring* of preverbal speech and efficient *encoding* are necessary to maintain fluent, accurate speech production. At least 3 contemporary theories of stuttering implicate atypical preverbal monitoring as an area of compromise in persons who stutter [30–32]. If this is the case, differences attributed to phonological encoding in AWS observed in previous studies, and by extension differences in everyday speech production, may have been confounded by atypical self-monitoring abilities. The present study examined whether preverbal monitoring differences may have contributed to the outcomes of one experimental method – the silent phoneme monitoring task – that has been used to support phonological encoding insufficiency in AWS. One method of measuring monitor function is assessment of uniqueness point effects during silent monitoring. Ozdemir et al. [33] found that AWNS monitor phonemes located after the uniqueness point of a word – the first phoneme that isolates the word from all lexical competitors – faster than phonemes located before this critical phoneme. The uniqueness point effect was interpreted by these authors to reflect a perceptual-based phenomenon observed when monitoring both internal and external speech and therefore linked to the function of the preverbal monitor. Thus, to examine the function of the preverbal monitor in AWS relative to AWNS, we examined whether the uniqueness point effect was also observed for AWS when silently monitoring internal speech.

Phonological Encoding and Monitoring in Typically Fluent Adults

As proposed by Levelt and colleagues' [29] model of speech production, speech planning requires 3 distinct levels of processing prior to production. The first level – the conceptualizer – determines communicative intent and activates the corresponding semantic and lexical representations, also known as the lexeme (e.g., [escort] + tense + aspect). The second level of processing – the formulator – is comprised of 2 sublevels of processing: grammatical encoding and phonological encoding. The grammatical encoding system accesses the phonological form as well as the appropriate morphological form (e.g., "escort*ing*"), known as the lemma, which serves as input to the phonological encoding system. The phonological system then activates the corresponding segmental properties and the metrical properties. Segmental properties include all corresponding phonemes (i.e., [ə s k o r t ŋ]) within the lemma, which are activated simultaneously and without syllable position assignment. Metrical properties include the number of syllables (indicated by _), syllable boundaries (indicated by /), and stress assignment (indicated by ^{1}) (i.e., ___ / ^{1} / ___) within the lemma. Once segmental and metrical properties are activated, individual sounds are assigned to each syllabic frame starting with the initial phoneme (i.e., /ə/) and ending with final phoneme (i.e., /ŋ/) to create 3 phonological syllables (i.e., [ə-^{1}skor-tŋ]). The rapid recombination of segmental and metrical properties is referred to as syllabification. The syllabified word form serves as the input for the third system – the articulator – wherein motor planning and execution is initiated. The process of speech production, from concept preparation to the onset of speech-motor preparation, is thought to occur within a rapid, finite time frame (e.g., ~600 ms per word [34]).

An equally important feature of Levelt and colleagues' [29] model of speech production that operates within this brief time frame is the preverbal speech monitor (i.e., the "perceptual loop" [35]). The purpose of the internal monitor is to examine the phonemic integrity of the speech plan and correct potential encoding errors just before articulation. In the model of Levelt et al. [29], the preverbal speech monitor is based in the conceptualizer and, unlike other levels of processing, can be consciously controlled by the speaker. Researchers have capitalized on speakers' ability to silently monitor the preverbal speech plan to estimate the time course of phonological encoding using silent phoneme monitoring tasks (e.g., [36, 37]). Silent phoneme monitoring tasks were derived from perceptual phoneme monitoring tasks designed to examine speech comprehension (for review, see [38]). During these tasks, participants were instructed to identify target phonemes within auditory targets as quickly as possible. Similarly, participants in silent phoneme monitoring tasks were instructed to identify target phonemes (e.g., /g/) in a target word (e.g., "*magnet*") as quickly as possible via button press, but in the absence of auditory input or speech production (see [18] for discussion of cueing techniques). Data from silent phoneme monitoring tasks found that adults identify phonemes within a target word at incrementally slower latencies, from beginning to end, prior to production (e.g., /m/</g/</n/</t/), supporting Levelt and colleagues' prediction that phonemic sequences are generated in a left-to-right manner prior to production in typically fluent speakers. If moments of stuttered speech occur in AWS, at least in part, due to delayed or less efficient phonological encoding processes prior to execution [3, 4], dissimilar response latencies during silent phoneme monitoring would be expected due to slower or less efficient assembly of upcoming speech plans.

Silent Phoneme Monitoring in AWS

A number of studies have compared silent phoneme monitoring abilities between AWS and AWNS. Sasisekaran et al. [24] found AWS (n = 10) monitored C_1–C_4 phonemes within $C_1VC_2C_3VC_4$ stimuli (e.g., *mattress*, *napkin*) significantly slower than AWNS (n = 11). Sasisekaran and de Nil [39] found AWS (n = 10) monitored syllable-offset phonemes (i.e., C_2, C_4) within C_1VC_2-C_3VC_4 compound words (e.g., "*hotdog*," trochaic stress) and noun phrases (e.g., "*hot dog*," iambic stress) more slowly than AWNS (n = 12). Sasisekaran et al. [25] found adolescents who stutter (n = 9; 10 to 14 years of age) were slower to monitor C_1–C_4 consonants within C_1VC_2-C_3VC_4 stimuli (e.g., "*basket*") compared to nonstuttering age-matched peers (n = 9). Coalson and Byrd [23] instructed AWS (n = 22) and AWNS (n = 22) to silently monitor phonemes within $C_1VC_2C_3VC_4$ nonwords with either trochaic stress (e.g., "*FAZmool*") or iambic stress (e.g., "*fazMOOL*") but identical phonemic composition. Although no between-group differences in monitoring latencies were observed, within-word monitoring patterns between C_2 and C_3 were delayed within AWS for stimuli with less frequent iambic stress. As suggested by Sasisekaran [40], these collective differences in monitoring performance by participants who stutter indicate potential weaknesses in the incremental, time-constrained process of phonological encoding, due to inefficiencies in segmental encoding and/or metrical encoding. However, due to reliance on the preverbal monitor during silent phoneme monitoring tasks, it is possible that these differences may also reflect deficits in the preverbal monitoring system itself. This possible confound has not been directly explored, but may help to reconcile (a) differences observed for AWS during silent monitoring tasks, and (b) the relatively null findings observed between AWS and AWNS during tasks that do not rely on preverbal monitoring abilities (i.e., implicit priming [26, 27]; auditory priming [28]).

Speech Monitoring in AWS

Three models of stuttering implicate potential deficits in preverbal speech monitoring in persons who stutter rather than, or in addition to, deficits in speech formulation. Vasic and Wijnen's [32] Vicious Circle Hypothesis (VCH) modified Postma and Kolk's [4] Covert Repair Hypothesis and proposes that individuals who stutter do not exhibit encoding deficits, but instead develop a hypervigilant speech monitor that identifies subtle temporal or rhythmic variation in segmental units during formulation as phonological errors. This overdetection of nonexistent errors then leads to overcorrection of accurate speech plans which, in turn, compromises speech fluency. The Variable Release Threshold (VRT) hypothesis proposed by Brocklehurst et al. [31] modified Howell's [3] EXPLAN model to suggest that during moments of stuttered speech, the threshold of acceptable phonemic quality to release an intended speech plan is elevated to excessive levels to accommodate real (or anticipated) imperfections, thereby delaying or halting preparation of speech plans and subsequent execution. Arenas's [30] Speech and Monitoring Interaction (SAMI) framework expands the notion of excessive or hypervigilant monitoring proposed by the VCH and VRT, but acknowledges that latent difficulties during speech formulation may also be present in persons who stutter. The SAMI model proposes that increased contextual demands may aggravate a vulnerable preverbal monitor in individuals who stutter and, as a result, trigger compensatory cognitive processes (e.g., increased attention, motor inhibition) in response to potential difficulties formulating speech. Across these theories, maladaptive or hypervigilant preverbal monitoring delays the completion of speech plans prior to production, which results in halted or repetitive articulatory responses that are characteristic of stuttered speech.

These 3 theories describe the preverbal monitor as a domain-general process rather than perception-based system proposed by Levelt et al. [29]. Nevertheless, all 3 theories cite evidence for overly critical evaluation of internal *and* external speech production by AWS. For example, Lickley et al. [41] found AWS are more likely to judge external speech as less fluent regardless of whether the speaker is an AWS or a typically fluent adult. Although these participants provided subjective ratings of "fluency," the authors interpreted AWS's propensity to judge third-party speech that was ostensibly fluent more harshly than AWNS as evidence for the VCH [32]. Lickley et al. [41] attributed this pattern in AWS to (a) small vocal deviations that may be present in the fluent speech of AWS compared to the fluent speech of AWNS, and (b) a monitoring system in AWS that was more sensitive to these subtle differences. Differences in perceptual evaluation of phonological code in AWS are not restricted to third-party spoken word production. Arnstein et al. [42] found stronger activity in AWS in neural regions associated with error detection during a silent rhyme judgment task compared to AWNS, even in the absence of actual error, suggesting the preverbal monitor in AWS may be prone to overly rigid evaluation, similar to their external monitor. In the same study, error detection differences between AWNS and AWS approached but did not reach significance during a nonlinguistic flanker task, wherein participants must quickly and accurately judge congruence of 2 sets of arrows presented simultaneously. This dissociation suggests that AWS may exhibit unique difficulties only when monitoring preverbal phonological information rather than a general monitoring deficit. These data coincide with Sasisekaran and colleagues [24, 25] who found no differences between persons who do and do not stutter when monitoring a series of nonlinguistic pure tones, suggesting minimal differences in the perception of nonlinguistic acoustic signal.

Data to support preverbal monitoring deficits in AWS due to aberrations in the perceptual system are not without exception. Although group differences when monitoring phonemes within inner speech were detected by Sasisekaran and de Nil [39], no significant latency or accuracy differences for AWS were detected when monitoring C_2 and C_4 positions of the same auditory stimuli. Further, Postma and Kolk [43] found AWS were equally accurate when identifying their speech errors under auditory masking as AWNS, suggesting that any self-monitoring differences in AWS are not perception-based, but instead associated with a phonological processing deficit (see also [44]). In this same study, however, and in contrast to Lickley et al. [41], AWS detected significantly *fewer* errors in third-party speech than AWNS, suggesting that AWS may have simply generated fewer phonological errors than AWNS. Thus, available data indicate that error monitoring in AWS may be atypical compared to AWNS (cf. [43, 44]), but it remains unclear to what degree this difference is perceptual in nature, and to what degree this negatively impacts preverbal monitoring during speech planning in AWS.

Uniqueness Point Effects

Whether or not deficits in perceptual monitoring may influence internal monitoring in AWS can be further investigated by examining whether well-established patterns observed during perceptual monitoring tasks are also observed when monitoring internal speech plans. One pattern, in particular, that can be examined is the uniqueness point effect [38]. The *uniqueness point* of a word is defined as the sound, from beginning to end of the word, where it diverges from all other morphologically unrelated words (e.g., [45, 46]). Uniqueness point effects have been observed during perceptual phoneme monitoring studies in adults to assess the influence of lexical competitor environment upon word recognition. Marslen-Wilson [45, 46] found that adult listeners were able to rapidly distinguish acoustically presented words from nonwords before hearing the entire word and immediately after hearing the critical phoneme (e.g., /k/ in "*trenker*") that eliminated all known words which shared the same initial phonetic sequence (e.g., "*<u>tren</u>ch*," "*<u>tren</u>d*"). Gaskell and Marslen-Wislon [47] found repetition priming in adults was fastest when the related phonological primes (e.g., *gar<u>m</u>*-) included the uniqueness point of the word (e.g., *garment*), and faster still when the uniqueness point occurred earlier in the phonological target (e.g., approximately the third phoneme in 6-phoneme targets). Frauenfelder et al. [48] examined the contribution of lexical information during perceptual phoneme monitoring and found that identification of phonemes located after the uniqueness point occurred more rapidly than for phonemes located before the uniqueness point. These findings suggest that phonological information of a target word becomes immediately available once the identity of the lexical item is unambiguous (i.e., few or no matching competitors). These data support top-down activation of sublexical information upon isolation and full activation of a single lexical item [cf. 49]. However, these studies were designed to measure lexical access during spoken word recognition, rather than the efficiency of the perceptual monitor. Thus, these data were not designed to verify the claim of Levet et al. [29] that same speech comprehension system used to monitor external speech also supported internal speech monitoring.

To more directly assess the claim of Levelt et al. [29] that the comprehension system governs the preverbal speech monitor, Ozdemir et al. [33] examined whether uniqueness point effects were also observed during a silent phoneme monitoring task. In their study, Dutch-speaking adults ($n = 32$) were instructed to press a button as quickly as possible if the name of the pictured object contained the target phoneme. Similar to perceptual phoneme monitoring [48], participants were faster to identify sounds if they occurred *after* the uniqueness point of the word (when competing words have been eliminated) compared to sounds *before* the uniqueness point (when competing words remain candidates for selection). Moreover, the greater the distance of a target phoneme after the uniqueness point, the more rapidly the final phonemes were identified by the speaker. That is, once the uniqueness point of a target word had been reached during phonological encoding (e.g., /m/ in "*nutmeg*," or C_3), the remaining phonological content (e.g., /-ɛg/, or -VC_4) became immediately available for motor processing. In contrast, participants did not exhibit differences in picture-naming latencies regardless of uniqueness point position. Authors concluded that because perception-specific effects observed during auditory phoneme monitoring were observed during silent phoneme monitoring, but not speech production, the preverbal speech monitor is primarily based in the speech comprehension system.

Evidence that top-down lexical mediation may influence the response latency of the preverbal monitor in AWNS in a predictable manner that is theoretically independent from phonological encoding provides an opportunity to assess whether previous differences noted in AWS performance in phoneme monitoring tasks may be confounded by preverbal monitoring deficits. For example, the presence of a uniqueness point 1 or 2 phonemes before word-final phonemes may have accelerated identification of word-final sounds (i.e., C_4) for AWNS in previous phoneme-monitoring studies in a manner similar to Ozdemir et al. [33]. If preverbal monitoring deficits were present in AWS, however, participants may have been slower to detect, or benefit from, uniqueness points and, consequently, may have identified C_4 targets more slowly than AWNS. Examination of word-final phoneme monitoring latencies between AWS and AWNS in previous silent phoneme monitoring tasks [24, 25] suggests this may be the case.

Specifically, Sasisekaran et al. [25] reported that adolescents who stutter exhibited longer C_4 monitoring latencies compared to C_3 (visual inspection of Figure 1 of [25], approximately +250 ms), whereas nonstuttering peers exhibited smaller C_3–C_4 latency differences (visual inspection of Figure 1 of [25], approximately +100 ms). In a study that compared mature speakers who do and do not stutter (18–49 years of age), Sasisekaran et al. [24] reported slower (but nonsignificant) monitoring of C_4 phonemes than C_3 phonemes for AWS (mean = +67.9 ms, SD = 178.8) compared to a nonsignificant *decrease* observed in AWNS (mean = –25.4 ms, SD = 113.7 ms). Findings of overall "levelling-off" of latency responses in AWNS is consistent with previous studies by Wheeldon and colleagues [36, 37] who also reported small and nonsignificant C_3–C_4 latency differences for typically fluent adults (mean = 13 and 25 ms, respectively). Deviation from this expected within-word latency monitoring patterns in AWS suggests deficiencies in the preverbal monitor may have also contributed to the atypical responses latencies. However, the uniqueness point of stimuli was not held constant in these studies, and between-group monitoring differences for word-final phonemes cannot be attributed to this factor alone. To more directly test this assumption, AWNS and AWS would need to silently monitor phonemes within stimuli for which the uniqueness point has been systematically varied.

Rationale and Purpose of the Present Study

In sum, previous studies found that AWS exhibit slower or atypical identification of word-final target sounds than AWNS during silent phoneme monitoring tasks. These differences are thought to reflect inefficiencies in the phonological encoding system in AWS. It is also possible differences in silent phoneme monitoring may reflect perception-based deficits in the preverbal monitor rather than, or in addition to, production-based phonological encoding deficits. If this is the case, one would expect perception-specific effects observed in AWNS during silent phoneme monitoring task, such as uniqueness point effect [33], to be less robust or absent in AWS during such tasks. To assess this prediction, AWNS and AWS in the present study completed a silent phoneme monitoring task comprised of stimuli that either had an early uniqueness point (EUP), late uniqueness point (LUP), or no uniqueness point (NUP). We hypothesized that if the internal speech monitor is similar in AWS and AWNS, word-final phonemes would be identified more rapidly for words with EUP and LUP than NUP due to rapid isolation and access to the word within the lexicon by both groups. If, however, AWS are comparable in their identification of word-final phonemes within words with EUP, LUP, and NUP, findings would support atypical operation of the internal speech monitor in AWS during these tasks and, by extension, during typical speech production.

Methods

Participants

Twenty adults (10 AWNS, age range = 19–22 years, mean = 20.9, SD = 0.88; 10 AWS, age range = 19–54 years, mean = 28.6, SD = 12.16) completed 2 separate 90-min sessions. The first session consisted of (a) intake questionnaire including general demographics, current medication use, medical history, and previous speech-language treatment, (b) 5 phonological processing subtests (Comprehensive Test of Phonological Processing [CTOPP; 50]; Comprehensive Test of Phonological Processing – 2nd Edition [CTOPP-2; 51]), (c) vision and hearing screening [52], (d) handedness inventory [53], and (e) language profile questionnaire (Language History Questionnaire [LHQ; 54]). The second session included the familiarization task, silent picture recognition task, and silent phoneme monitoring task described by Ozdemir et al. [33]. Participants were excluded from the study if he or she did not pass the hearing screening, currently used psychotropic medications, or presented with additional medical or speech-language concerns that would influence completion of the experimental task. All participants included in the study were actively completing or had completed a baccalaureate degree. Each participant was provided their informed consent during the first session, and the study protocol was approved by the institute's committee on human research.

Talker Classification and Stuttering Severity

Participants were classified as an AWS if he or she self-identified as a person who stutters and had previously received a diagnosis of stuttering by a licensed speech-language pathologist. Participants were required to receive a score of 10 or higher on the Stuttering Severity Instrument – 4th Edition (SSI-4; [55]) to be considered an AWS. Intrarater reliability was determined based on SSI-4 score for all 20 conversational speech samples by the first author (90.0% agreement; kappa = 0.85), and interrater reliability was established for all 20 samples between the first author and a graduate research assistant trained in disfluency assessment (85.0% agreement; kappa = 0.77). See Table 1 for detailed breakdown of stuttering frequency and severity per participant. To account for any potential influence of stuttering severity and experimental outcomes, individual scores from the SSI-4 were included as a covariate in post hoc analysis.

Language History Profile

The stimuli used in the present study were based on phonetic sequences within the English language. In theory, knowledge of additional languages may alter the uniqueness point of individual stimuli if a speaker had multiple lexical entries from other languages. Therefore, data from the LHQ [54] were collected as a screening measure to ensure each participant was a native English speaker with minimal knowledge of additional languages. Self-rated proficiency in English and any language other than English as reported for Item 12 on the LHQ (i.e., 7-point Likert scale, averaged across modalities; 1 = Very poor, 7 = Native) is provided for each participant in Table 1. Of the 10 respondents who reported any knowledge of another language, 5 participants self-rated proficiency between "Very poor (1)" and "Poor (2)," 2 participants self-rated between "Poor (2)" and "Fair (3)," and 3 participants self-rated proficiency between "Functional (4)" and "Good (5)." These 10 participants also reported second language current use as minimal (less than 15 min per month on average per week; Item 15), and that second languages were first learned no earlier than 14 years of age (Item 9). Similar to stuttering severity, individual proficiency scores from the LHQ were included as a covariate in post hoc analysis to account for any potential influence upon experimental measures.

Phonological Processing Subtests

The purpose of the preexperimental phonological subtests was to ensure no participant presented with frank, underlying deficits in phonological encoding or phonological working memory that may influence experimental outcomes irrespective of factors manipulated for the purpose of the study. Subtests included Nonword Repetition (CTOPP-2-Subtest V [51]), Word Segmentation (CTOPP-Subtest XI [50]), Nonword Segmentation (CTOPP-2-Subtest IX), Forward Digit Span (CTOPP-2-Subtest IV), and Backward Digit Span (CTOPP-2-Subtest IV, modified). These subtests were selected to provide a cross-section of data that distinguished phonological awareness abilities from phonological working memory abilities. Table 1 depicts raw scores for all participants across each subtest. A series of independent *t*-tests revealed no significant differences in phonological processing abilities between groups (Nonword Repetition: $p = 0.303$; Word Segmentation: $p = 0.773$; Nonword Segmentation: $p = 0.314$; Forward Digit Span: $p = 0.514$; Backward Digit Span: $p = 0.944$).

Stimuli Development

Stimuli used in the present study during silent phoneme monitoring were comprised of 3 types of words: words with an EUP, words with a LUP, and words with NUP. Each condition contained 5 words, for a total of 15 target words per participant. An initial cohort of 177 words were compiled from 3 sources: (a) 42 EUP and 42 LUP auditory stimuli included in Gaskell and Marslen-Wilson [47] (Appendix A – 5–7 phonemes in length), (b) 10 picture stimuli used in previous silent phoneme monitoring tasks ([24]; 6 phonemes in length), and (c) 83 additional words generated using the English Lexicon Project (ELP) online database (http://elexicon.wustl.edu/ [56]; 5–7 phonemes in length). ELP search was restricted to bisyllabic, monomorphemic content words. Each of the 177 potential words was coded for uniqueness point position, and 6 additional phonological and lexical factors (i.e., word frequency, neighborhood density, neighborhood frequency, segmental phonotactic probability, biphone phonotactic probability, and word familiarity).

Uniqueness Point

Uniqueness point for each word was determined using the searchable phonetic database of English produced with General American dialect within the *Longman Pronunciation Dictionary*, ed 3 [57]. The accompanying Sound Search feature enables users to search for words that start or end with a particular sound or sound sequence. In this system, the asterisk [*] symbol can be entered into the search query to return any number of phonemes at

Table 1. Participant demographics, stuttering severity, and performance on phonological processing tasks

	Age, years	Gender	Eth	%SS	SSI-4[a]	Severity[a]	Self-ID	Prev Dx	NWR[b]	WS[c]	NWS[b]	FDS[b]	BDS[b]	Eng Prof[d]	Non-Eng Prof[d]
AWNS-1	22	F	C	0.67	5	None	N	N	17	14	25	19	16	7.0	4.5[F]
AWNS-2	21	F	C	0.33	5	None	N	N	22	17	30	25	15	7.0	None
AWNS-3	22	F	C	0.33	5	None	N	N	21	16	29	24	18	7.0	5.25[S]
AWNS-4	21	F	C	0.67	6	None	N	N	19	15	23	27	14	7.0	None
AWNS-5	20	F	C	0.33	7	None	N	N	18	16	27	17	12	7.0	None
AWNS-6	21	F	AA	0.67	6	None	N	N	19	14	27	19	16	7.0	4.0[S]
AWNS-7	21	F	C	0.67	6	None	N	N	19	10	23	20	11	7.0	None
AWNS-8	19	M	C	1.00	6	None	N	N	18	18	27	21	15	7.0	None
AWNS-9	21	M	C	0.67	6	None	N	N	20	15	23	24	18	7.0	None
AWNS-10	21	M	C	1.67	7	None	N	N	18	21	12	12	20	7.0	2.5[S]
AWS-1	21	M	C	8.00	17	VM	Y	Y	18	14	23	19	13	7.0	None
AWS-2	19	M	C	39.00	35	Sev	Y	Y	20	18	25	20	15	7.0	None
AWS-3	45	M	AA	3.00	10	VM	Y	N	19	15	27	23	21	7.0	1.5[F]/1.5[S]
AWS-4	22	M	C	14.33	26	Mod	Y	Y	20	20	27	21	19	7.0	1.5[S]
AWS-5	54	M	C	6.67	11	VM	Y	Y	14	12	23	19	12	7.0	None
AWS-6	26	F	AA	3.33	10	VM	Y	Y	16	14	30	18	12	7.0	1.0[P]
AWS-7	19	M	C	5.00	11	VM	Y	N	19	17	30	22	13	7.0	3.0[S]
AWS-8	34	M	C	12.00	32	Sev	Y	Y	21	18	30	18	19	7.0	1.0[G]/1.0[S]
AWS-9	19	M	C	5.67	19	Mild	Y	Y	16	9	24	14	12	7.0	1.0[C]
AWS-10	27	M	C	7.00	22	Mild	Y	Y	19	15	26	23	18	7.0	None

AWS, adult who stutters; AWNS, adult who does not stutter; C, Caucasian; Eth, ethnicity; AA, African-American; %SS, percent stuttered-like disfluencies per 100 words [60]; severity, none: no stuttering; VM, very mild; Mod, moderate; Sev, severe; VS, very severe; self-ID, self-identification as a person who stutters; Prev D, diagnosis of stuttering prior to 7 years of age. Phonological processing subtest scores reported in raw score to accommodate participants older than normative data. NWR, nonword repetition (subtest V – ceiling 30); WS, word segmentation (subtest XI – ceiling 20); NWS, nonword segmentation (subtest IX – ceiling 31); FDS, forward digit span (subtest IV – ceiling 28); BDS, backward digit span (subtest IV; ceiling – 28); Eng Prof, self-rated proficiency in English in terms of speaking, comprehension, reading, and writing averaged across 4 7-point Likert scales (1 = very poor, 7 = native); Non-Eng Prof, self-rated proficiency in any language other than English in terms of speaking, comprehension, reading, and writing averaged across 4 7-point Likert scales (1 = very poor, 7 = native, none = no reported knowledge of other languages); [S] Spanish; [F] French; [P] Portuguese; [G] German; [C] Chinese. [a] SSI-4 [55]; [b] CTOPP-2 [51]; [c] CTOPP [50]; [d] LHQ – Item 12 [54].

that serial position. For example, [*ɪs] generates a list of all phonetic words ending with /-ɪs/ (e.g., *abyss*, *avarice*, *actress*…), and [ret*] generates a list of all phonetic words starting with /ret-/ (e.g., *retina*, *retribution*, *wretch*…). Target words in the present study were isolated from all other competing words by entering each phoneme individually into the Sound Search function, from beginning to end, followed by the asterisk symbol. The uniqueness point position of each word was established if inclusion of the asterisk returned no additional, unrelated words. Similar to previous studies of uniqueness point [45–48], morphological variants, hyphenated words, proper nouns, homophones, homographs, and extremely low frequency words were excluded as they are not considered to be unique or substantial lexical competitors (see Appendix for remaining neighbors at established uniqueness point and, if available, associated word frequency values).

Lexical and Phonological Properties

Six additional factors known to influence phonological processing and speed of response in AWNS were calculated for each word and included as covariates during analysis: word frequency, neighborhood density, neighborhood frequency, segmental phonotactic probability, biphone phonotactic probability, and word familiarity. Word frequency, neighborhood density and frequency, and segmental and biphone phonotactic probability values were determined using the online CLEARPOND linguistic database at http://clearpond.northwestern.edu/index.html [58]. Phonological and lexical values provided in the CLEARPOND database were derived from approximately 51 million words included in subtitle corpora taken from 8,388 US films and television series (SUBTLEX-US [59]). Word familiarity values were obtained from the MRC database developed by Coltheart ([60]; http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.html).

Stimuli Selection

Of the initial 177 candidate words, 38 were excluded because description of phonetic length was inconsistent across 3 phonetic databases (Longman Pronunciation Dictionary [57], CLEARPOND [58], MRC [60]). Of the remaining 139 words, 81 were excluded based on difficultly of visual presentation (e.g., *nephew*, *her-*

Table 2. Target stimuli and corresponding uniqueness point, phonological properties, and lexical properties

	IPA[a]	UP[a]	Target	Word shape	WF[b]	ND[b]	NF[b]	PP-S[b]	PP-B[b]	Fam[c]
Early uniqueness point										
Dolphin	/d ɑ l f ɪ n/	/f/	/n/	CVCCVC	2.76	1	3.96	0.34	0.02	0
Falcon	/f æ l k ə n/	/k/	/n/	CVCCVC	3.31	0	0	0.37	0.05	0
Napkin	/n æ p k ɪ n/	/k/	/n/	CVCCVC	3.61	1	2.57	0.32	0.02	495
Tadpole	/t æ d p o l/	/p/	/l/	CVCCVC	0.59	1	0.55	0.27	0.01	0
Walrus	/ w ɑ l r ə s /	/r/	/s/	CVCCVC	1.12	0	0	0.31	0.03	506
Late uniqueness point										
Biscuit	/b ɪ s k ɪ t/	/ɪ/	/t/	CVCCVC	3.75	1	3.27	0.41	0.03	521
Closet	/k l ɑ z ə t/	/ə/	/t/	CCVCVC	27.08	1	2.14	0.31	0.02	540
Forest	/f o r ə s t/	/s/	/t/	CVCVCC	18.88	2	2.00	0.33	0.04	513
Spinach	/s p ɪ n ə ʧ/	/ɪ/	/ʧ/	CCVCVC	2.55	0	0	0.25	0.02	452
Stomach	/s t ^ m ə k/	/ə/	/k/	CCVCVC	33.82	1	1.12	0.27	0.03	547
No uniqueness point										
Bandage	/b æ n d ɪ ʤ/	/ʤ/	/ʤ/	CVCCVC	2.86	2	1.11	0.33	0.05	546
Compass	/k ^ m p ə s/	/s/	/s/	CVCCVC	4.06	2	5.69	0.34	0.04	0
Magnet	/m æ g n ə t/	/t/	/t/	CVCCVC	2.75	1	0.90	0.32	0.02	526
Mattress	/m æ t r ə s/	/s/	/s/	CVCCVC	6.61	0	0	0.35	0.04	524
Mustard	/m ^ s t ɚ d/	/d/	/d/	CVCCVC	6.45	3	1.37	0.37	0.05	532

IPA, transcription in International Phonetic Alphabet; UP, uniqueness point; WF, word frequency per million words; ND, neighborhood density; NF, neighborhood frequency; PP-S, segmental phonotactic probability; PP-B, biphone phonotactic probability; Fam, word familiarity. [a] *Longman Pronunciation Dictionary*, ed 3 [57]. [b] CLEARPOND lexical database [58]. [c] MRC database [60].

mit, *gospel*). Of the remaining 58 words, items were excluded if the word contained repeated consonants ($n = 9$) or lexical-phonological properties were not available in the CLEARPOND database ($n = 13$), resulting in 36 words that were either 5 phonemes ($n = 14$) or 6 phonemes ($n = 22$) in length. Six-segment targets were selected to maximize stimuli and maintain consistent distance across EUP, LUP, and NUP conditions. Of these 22 words, 3 were excluded as compound words which arguably contained more than 1 free morpheme (i.e., *doorknob*, *necklace, sailboat*). A total of 19 target word candidates remained for the experimental task (6 EUP, 6 LAP, 7 NUP).

Pilot testing conducted prior to experimental testing ($n = 6$ AWNS) was conducted to assess verbal accuracy of the 19 items during a picture naming task. Four items with the lowest accurate identification were removed (e.g., *volume*, *captain*, *bandit*, *public*), resulting in 15 total words used during the experimental portion of the task (5 EUP, 5 LUP, 5 NUP). See Table 2 for target stimuli within each condition and associated lexical and phonological properties.

Stimuli Recording

All auditory stimuli were generated by a monolingual English-speaking female with no reported history of speech, language, or hearing difficulties. Digital audio files were collected using KayPENTAX Computerized Speech Lab (KayPENTAX, Lincoln Park, NJ, USA) files at a 22,050-kHz sampling rate and 16-bit quantization.

Picture Stimuli

Due to the stringent criteria necessary during stimuli development, selection of corresponding visual stimuli to represent target words using established picture stimuli corpora was a challenge. Picture stimuli for target words were initially searched via the 520 object pictures provided in the International Picture Naming Project database [61]. However, only 3 of the 15 target words were available within this large database. Instead, 4.5” × 4.5” color images available via online image search were selected and used during experimental portions of the study. Similar to Ozdemir et al. [33], item analysis of picture recognition latencies was included in the final experimental study to confirm whether recognition of individual pictures differed between items overall within and across the 2 primary variables: talker group (AWNS, AWS) and uniqueness point (EUP, LUP, NUP).

Procedure

The experimental silent phoneme monitoring task in the present study was based on the paradigm described by Ozdemir et al. [33]. Participants completed 4 experimental tasks during the second session: phoneme exposure, familiarization, silent picture recognition task, and silent phoneme monitoring task. Participants were seated comfortably in front of a computer inside a sound booth. All auditory stimuli were presented at 45 dB SPL via headphones. Manual button-press latencies during the picture recognition and silent phoneme monitoring task were collected via SuperLab Pro (v 4.5) stimulus presentation software and corresponding

RB-740 keypad. All post-trial verbal responses were recorded via audio-video camera (Q2 Zoom, Tokyo, Japan) and analyzed for phonemic accuracy and fluency.

Phoneme Exposure

Similar to Coalson and Byrd [23], participants completed a phoneme exposure task before the experiment was administered. Participants were instructed to first listen to each target phoneme [s, n, l, k, t, ʧ, ʤ] presented simultaneously via headphones and orthographic letter displayed on the monitor. Each phoneme was presented 4 times with no participant response required. Upon completion, participants were faced away from the monitor and asked to provide the corresponding letter for the phoneme presented via headphones. The phoneme exposure task was repeated until the participant identified each phoneme orally with 100% accuracy (28 of 28 opportunities).

Familiarization

Prior to each experimental block, participants completed a brief familiarization task of the upcoming visual stimuli. All 30 pictures within the block were presented simultaneously with target word in written form in randomized order. Participants were instructed to say each word aloud as it appeared. Trials were self-paced to accommodate for stuttered speech, if present. All participants were required to repeat all 30 words within the experimental block with 100% accuracy before proceeding to silent picture recognition.

Silent Picture Recognition

Similar to Ozdemir et al. [33], participants completed a silent picture recognition task prior to the experimental task. The purpose of the silent picture naming task was to assess potential differences in picture recognition upon manual response latencies. Participants read the following instructions prior to the task: “You will see a word, then see a picture. If the word and picture match, press ‘Yes’ as quickly as possible. If not, press ‘No.’ Press any button to begin.” Each trial began with an orienting cross presented in the middle of the screen (1 s), followed by a word in orthographic form (0.5 s), a second orienting cross (0.5 s), and then a picture displayed in the center of the screen. Participants were provided 1.5 s to press the “Yes” or “No” button after picture presentation, before proceeding to a prompt screen (“Ready?”) which triggered the next trial upon button press. Each target word within the block was presented twice, one true positive match and one false negative mismatch per word-picture pair, resulting in 60 total trials. Reaction time latencies were measured from the onset of the picture presentation to button press.

Silent Phoneme Monitoring

The experimental silent phoneme monitoring task in the present study was based on the uniqueness point paradigm described by Ozdemir et al. [33]. Stimuli were presented via 15 pictures with bisyllabic names with either an EUP (*n* = 5), LUP (*n* = 5), or NUP (*n* = 5). In the EUP distance condition, the picture name became unique at 2 phonemes prior to the final, target phoneme. In the LUP distance condition, the picture name contained a uniqueness point 1 phoneme prior to the final, target phoneme. In the NUP distance condition, the uniqueness point coincided with the final, target phoneme. In addition, 105 bisyllabic filler words were included (15 of which contained the target phoneme in word-initial position, 30 of which contained the target phoneme within the medial consonant cluster, 60 of which contained none of the target phonemes) to ensure response fidelity and prevent anticipation of specific phonemes in any specific position. The 120 target words were randomly assigned to 4 experimental blocks each comprised of 30 items (15 true positive responses, 15 true negative responses per block). Presentation order of the 4 experimental blocks were counterbalanced between participants using Latin-square assignment prior to the session. This resulted in a total of 300 viable tokens (20 participants × 15 true positive responses) and a total of 100 words per uniqueness point condition (EUP, *n* = 100; LUP, *n* = 100; NUP, *n* = 100).

Each participant received 8 practice trials during their first experimental block and prior to the experimental silent phoneme monitoring task. These practice trials were identical to the silent phoneme monitoring experiment, with the exception that none of the picture stimuli included any of the 15 critical uniqueness point targets. Upon completion of practice trials, the participant began the silent phoneme monitoring task. Participants read the following written instructions before each block: “You will hear a sound, then see a picture. Press ‘Yes’ as quickly as possible if the sound is in the name of the picture. If not, press ‘No.’ Do not say the word aloud. Press any button to begin.” Each of the following 30 trials were presented in randomized order and were presented in the following sequence: “Ready?” screen (advanced on button press), orienting cross located in the center of the screen (2.5 s), simultaneous presentation of audio phoneme (e.g., /s/) with corresponding letter (e.g., “S”) on the screen (0.5 s), a second orienting cross in the center of the screen (1.5 s), followed by visual presentation of a target picture, which advanced upon button press or prespecified time frame (2 s). After participant had provided a manual response, or 2 s had elapsed, the participants were then prompted by on-screen instructions to “Say the word aloud.” Again, this final screen was self-paced to accommodate potential stuttered speech. Participants advanced to the beginning of the next trial by pressing any button. Upon completion of all 30 trials, the next experimental block began exactly as described but with a different stimuli set and without the preceding practice trials.

Data Coding and Reliability

To ensure comparison of uniqueness point effects during speech planning between groups was based on fluent, accurate generation of inner speech, responses that reflected inaccurate or atypical processing were eliminated from reaction time analysis. Therefore, manual response tokens were eliminated based on the following criteria:

(a) False negative response: target phonemes present in word-final position were identified as absent by participant

(b) No manual response: participant did not press Yes or No button during trial within the allotted response time (2 s)

(c) No verbal response: participant did not provide a verbal response during post-trial production

(d) Lexical error: participant identified the target word incorrectly during post-trial production

(e) Typical disfluency: post-trial verbal response contained an interjection or revision

(f) Stuttering-like disfluency: post-trial verbal response contained a stuttering-like disfluency (part-word repetition, prolongation, or block [62])

Table 3. Tokens excluded from analysis

	AWNS				AWS				Total
	EUP	LUP	NUP	*n*	EUP	LUP	NUP	*n*	
Initial corpus	50	50	50	150	50	50	50	150	300
False negative error	7	0	2	9	5	2	5	12	21
No manual response	1	2	2	5	2	3	4	9	14
Lexical error	0	0	0	0	0	0	0	0	0
No verbal response	0	1	0	1	1	0	0	1	2
Typical disfluency	0	0	0	0	0	0	0	0	0
Stuttering-like disfluency	0	0	0	0	4	5	2	11	11
Outlier	3	1	3	7	0	5	1	6	13
Technical error	5	1	4	10	5	0	3	8	18
Total excluded	16	5	11	32	17	15	15	47	79
Usable *n*	34	45	39	118	33	35	35	103	221

AWS, adult who stutters; AWNS, adult who does not stutter; EUP, early uniqueness point; LUP, late uniqueness point; NUP, no uniqueness point.

(g) Outlier: manual reaction time was 2 SD above or below the participant's mean reaction time during silent phoneme monitoring task, or

(h) Technical error: manual response was not collected by stimulus software and/or post-trial verbal response could not be coded due to technical issues or sound quality.

To establish interrater reliability of post-trial production, 100% of the 300 tokens were coded offline by 2 graduate assistants trained in disfluency measures (98%; kappa = 0.96). Any discrepancies between coders were discussed between both research assistants and the first author, a certified speech-language pathologist, to reach majority consensus.

Excluded Tokens

A total of 300 manual responses were collected during experimental task (15 target words × 20 participants). From the initial 300 response tokens, 79 (26.3%) were excluded, resulting in 221 fluent, accurate tokens during final analysis (AWNS, *n* = 118; AWS, *n* = 103). See Table 3 for breakdown of token exclusion per group and uniqueness point condition. In addition, 17 tokens (5.7%) were removed from picture recognition analyses due to technical error (*n* = 16 [5.3%], AWNS, *n* = 8 [2.7%], AWS, *n* = 8 [2.7%]) or no manual response during trial (*n* = 1 [0.3%], AWNS, *n* = 0 [0.0%], AWS, *n* = 1 [0.3%]).

Analyses

The primary purpose of the present study was to assess the relationship between Talker Group (AWNS, AWS) and Distance (EUP, LUP, NUP) upon the silent identification of word-final phonemes. A generalized linear mixed model was conducted using SPSS (v. 24) to assess latency (in ms, continuous variable) of AWS and AWNS when accessing phonological information after the critical uniqueness point has been encountered. Talker Group and Distance served as fixed effects. Each participant provided multiple responses for each target condition (EUP, LUP, and NUP) for each target nonword; thus, participants served as a random factor with individual words as a repeated measure. Multiple pairwise planned comparisons were conducted for all main effects and interactions using Bonferroni adjusted *p* values.

To assess potential differences in picture recognition latencies between and across groups (AWNS, AWS), and between and across uniqueness point conditions (EUP, LUP, NUP), a silent picture identification task was administered prior to a silent phoneme monitoring task. An additional generalized linear mixed model analysis was conducted with participant as random effects, individual words (i.e., Word) as repeated measures, Talker Group and Word as the fixed effects, and manual reaction time latency as the dependent variable. Satterthwaite approximation was applied during both analyses to account for small sample size by providing larger and more conservative estimates of standard error.

Results

Silent Picture Recognition

No significant differences in reaction time latencies were detected for Word $F(14, 37) = 1.72$, $p = 0.093$, or Talker Group $F(1, 166) = 0.06$, $p = 0.811$. Interaction between Word and Talker Group was also nonsignificant $F(14, 37) = 0.44$, $p = 0.950$. Based on these analyses, the 15 target pictures did not significantly differ in ease of recognition, regardless of talker classification. See Table 4 for mean reaction time latencies and recognition accuracy for each target word within and across talker groups.

Uniqueness Point Effect

Figure 1 depicts the mean overall reaction time for AWNS and AWS for each Distance condition. No significant main effect of Talker Group $F(1, 11) = 0.32$, $p =$

Table 4. Mean reaction time latencies during silent picture recognition task

	AWNS	AWS	Total
Early uniqueness point			
Dolphin	549.22 (52.66)	585.33 (52.66)	567.28 (37.23)
Falcon	598.44 (115.22)	673.67 (115.22)	636.06 (81.47)
Napkin	536.00 (43.88)	508.56 (43.88)	522.28 (27.54)
Tadpole	531.67 (38.95)	543.89 (38.95)	537.78 (27.54)
Walrus	600.56 (69.88)	639.00 (69.88)	619.78 (50.94)
Total	551.25 (44.77)	606.33 (45.38)	578.79 (31.88)
Late uniqueness point			
Biscuit	551.50 (63.69)	528.90 (63.69)	540.20 (45.03)
Closet	682.20 (69.65)	689.50 (69.65)	685.85 (49.25)
Forest	616.90 (52.24)	486.20 (52.24)	551.55 (36.94)
Spinach	652.60 (115.74)	614.40 (115.74)	633.50 (81.84)
Stomach	688.20 (94.90)	674.90 (94.90)	681.50 (67.11)
Total	638.26 (43.46)	598.78 (43.46)	618.52 (30.73)
No uniqueness point			
Bandage	660.60 (61.82)	672.30 (61.82)	666.45 (43.71)
Compass	551.56 (113.79)	688.22 (113.79)	619.89 (80.46)
Magnet	593.40 (71.84)	549.70 (71.84)	571.55 (50.80)
Mattress	592.78 (48.59)	543.44 (48.59)	568.11 (34.36)
Mustard	541.33 (53.31)	449.44 (53.31)	495.39 (37.70)
Total	581.22 (44.22)	591.50 (44.33)	586.36 (31.31)

Mean reaction time latencies in milliseconds for each target word by condition. Standard error in parentheses. AWNS, adult who does not stutter; AWS, adult who stutters.

0.585, or Distance $F(2, 41) = 1.12$, $p = 0.337$, was detected. However, a significant interaction was detected between Talker Group and Distance $F(2, 90) = 6.94$, $p = 0.002$. Decomposition of the significant interaction revealed that AWNS identified word-final phonemes significantly faster for stimuli that contained an EUP (mean = 1,034.28 ms, SE = 65.04 ms) compared to LUP (mean = 1,153.01 ms, SE = 81.36 ms, $p = 0.037$) and NUP (mean = 1,135.31 ms, SE = 74.95 ms, $p = 0.043$). However, no significant latency differences were detected for AWNS between LUP and NUP conditions ($p = 0.584$). Overall, faster identification of C_4 targets by AWNS when silently monitoring words with EUP replicate the uniqueness point effects reported by Ozdemir et al. [33].

In contrast to AWNS, no significant differences in identification speed were detected for AWS between EUP (mean = 1,075.44 ms, SE = 60.96 ms) and NUP (mean = 1,093.71 ms, SE = 46.26 ms, $p = 0.771$), LUP and NUP (mean = 1,014.80 ms, SE = 51.51 ms, $p = 0.056$), or EUP and LUP ($p = 0.232$). That is, AWS identified word-final phonemes of internal speech with comparable speed re-

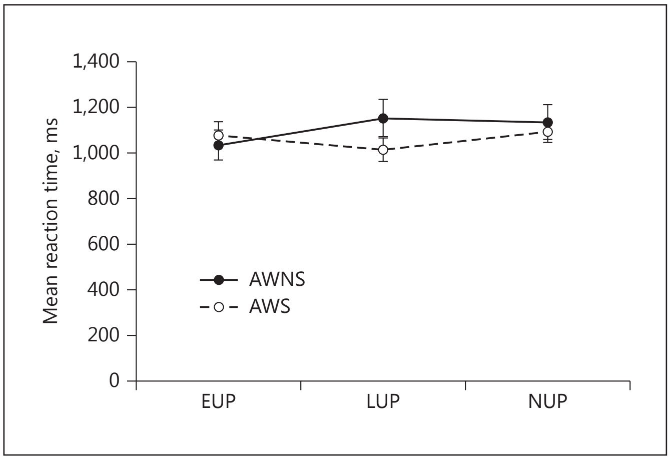

Fig. 1. Mean reaction time for adults who do and do not stutter (AWNS, AWS) silently monitoring word-final phonemes in words with an early uniqueness point (EUP; e.g., /dal**f**ɪn/), late uniqueness point (LUP; e.g., /bɪsk**ɪ**t/), and no uniqueness point (NUP; e.g., /mægnɛ**t**/). Word frequency, neighborhood density, neighborhood frequency, segment phonotactic probability, biphone phonotactic probability, and familiarity were included as covariates during analysis.

DOI: 10.1159/000485657

gardless of the presence or absence of uniqueness point, and indicate that, unlike AWNS, uniqueness point effects were not observed for AWS when monitoring internal speech.

In addition, 3 of the 6 covariates included in the model were also detected as significant: neighborhood density $F(1, 169) = 4.64$, $p = 0.033$, neighborhood frequency $F(1, 13) = 11.95$, $p = 0.004$, and familiarity $F(1, 14) = 6.94$, $p = 0.036$. Word frequency approached but did not reach significance $F(1, 135) = 2.93$, $p = 0.089$. Removal of covariates did not affect the significance of the Talker Group by Distance interaction $F(2, 120) = 7.15$, $p < 0.001$, the nonsignificant differences between Talker Group $F(1, 12) = 0.31$, $p = 0.587$, or the nonsignificant effect of Distance $F(2, 120) = 0.90$, $p = 0.409$, reported in the full model.

A final analysis was conducted to assess the potential influence of language profile and stuttering severity. To do so, SSI-4 score and cross-modal proficiency in any language other than English (Item 12, LHQ [54]) were included as covariates in addition to the 6 included in the original model. Inclusion of these covariates did not affect the significance of Talker Group by Distance interaction $F(2, 80) = 6.53$, $p = 0.002$, the nonsignificant differences between Talker Group $F(1, 5) = 1.26$, $p = 0.315$, or the nonsignificant effect of Distance $F(2, 37) = 1.14$, $p = 0.332$, reported in the original model. Neither stuttering severity $F(1, 1) = 1.54$, $p = 0.449$ nor proficiency in a second language $F(1, 7) = 0.20$, $p = 0.664$, reached significance.

Discussion

The purpose of the present study was to examine whether preverbal speech monitoring in AWS differed from AWNS. To do so, we examined whether AWS exhibited uniqueness point effects – a perception-based phenomenon that has been observed when AWNS monitor inner speech – during a silent phoneme monitoring task. AWS and AWNS identified, via button-press, the presence or absence of target phonemes in words that included either an EUP, LUP, or NUP. AWNS data replicated previous studies in nondisordered adults and confirmed that perception-based uniqueness point effects may be present when monitoring internal speech. The lack of uniqueness point effects in AWS, however, provides preliminary evidence to support differences in the preverbal monitoring of individuals who stutter.

Preverbal Speech Planning and Monitoring in AWS

Based on theoretical models that suggest atypical speech monitoring abilities in persons who stutter may contribute to moments of stuttering [30–32], we predicted AWS would not exhibit perception-specific uniqueness point effects exhibited by AWNS when monitoring inner speech [33]. This prediction was supported by significantly faster detection of phonemes that followed an EUP (but not LUP) for AWNS, but no phoneme monitoring latency differences for AWS irrespective of uniqueness point. These findings are consistent with previous studies that suggest atypical, perhaps overly rigid, evaluation of external speech ([41]; cf. [43]) *and* internal speech [42] by the perceptual system in AWS. Previous silent phoneme monitoring studies which indicate overall slower C_1–C_4 response latencies by AWS attributed these differences to phonological encoding difficulties [23–25]. Data from the present study provide an alternative perception-based account for these previous findings. That is, the speed and accuracy of syllabification in AWS may have been comparable to AWNS, but rapid and accurate detection of phonemes by the preverbal monitor in AWS may have been the reason for the additional time required to complete the task.

Perception-based theories do not negate the possibility that AWS were simply slower to encode phonological information in the present study and previous silent phoneme monitoring studies. However, monitoring deficits may explain the unexpected word-final (C_4) delays observed for individuals who stutter in Sasisekaran et al. [24, 25] that would not be predicted by encoding-based theories. Encoding-based accounts of stuttering such as the CRH [4] and EXPLAN [3] would predict that segment encoding deficits would occur most frequently at the typical loci of stuttered speech – at or near word onset (e.g., [63]). Faster C_4 identification in AWNS, but not AWS, may have occurred because the preverbal monitor in AWNS was able to quickly identify the uniqueness point and immediately accessed the remaining phonological segments. In AWS, however, an overly stringent monitor may have delayed identification of the critical phoneme and, instead, relied on incremental encoding for the remainder of the word. Reexamination of previous stimuli based on uniqueness point criteria used in the present study suggests this may be the case: at least 9 of the 14 stimuli in Sasisekaran et al. [24] and 6 of the 12 stimuli in Sasisekaran et al. [25] became unique at least 1 phoneme before the final, C_4 position. In both studies, AWS did not exhibit the expected "levelling-off" of latencies when monitoring C_3 and C_4 phonemes as observed for AWNS

[36, 37] or predicted by Levelt et al. [29], suggesting that the response to uniqueness points in stimuli used by previous studies may have been less robust in AWS compared to AWNS.

Diminished response to the uniqueness point within target stimuli due to atypical monitoring abilities may also provide a reasonable account for the prolonged C_3 and C_4 monitoring delays for younger children who stutter ([25], ages 10–14) compared to older AWS [24]. As noted by Frauenfelder et al. [48], the strength of uniqueness point effects would vary based on the participants' vocabulary size. Specifically, smaller vocabulary size would result in stronger uniqueness point effects because the target word has fewer lexical competitors to eliminate, and perhaps even shift the critical uniqueness point to an earlier position within the word. On the assumption that younger children have more restricted vocabularies than adults, one would expect stronger uniqueness point effect in nondisordered children and, by extension, C_4 phonemes to be identified as fast (or perhaps faster) than adults with larger vocabularies. For the nonstuttering adolescents in Sasisekaran et al. [25], this was the case – C_3 and C_4 phonemes were identified almost simultaneously with C_2 phonemes. Adolescents who stutter, however, required *significantly* longer to identify C_3 and C_4 phonemes than C_2 phonemes, suggesting that their preverbal monitor was less able to isolate the lexical item even amongst fewer lexical competitors. These claims should be taken with caution, as minimal evidence is available that reports whether speech planning processes in children are as susceptible to uniqueness point effects as adults (however, see Henderson et al. [64], for related effects in children 7–8 years old during auditory pause-detection task). Nonetheless, findings from the present study support further investigation of speech monitoring abilities in children who stutter using stimuli with systematically varied uniqueness points to assess if these preverbal monitoring differences are present at younger ages.

Theoretical Implications

The outcomes of the present study are consistent with theoretical models that propose that the preverbal monitor in persons who stutter may be hypervigilant (SAMI [30], VRT [31], VCH [32]). That is, if AWS apply an excessively high threshold for phonemic accuracy prior to motor execution, this may slow the detection of the critical uniqueness point phoneme. Slowed identification of the critical phoneme may, in turn, limit their ability to isolate it from competing lexical items and delay the availability of remaining phonological information. Further verification of each theoretical model would require specific modifications to assess the additional influence of cognitive factors (e.g., attention, anticipation, physiological arousal). For example, the VCH may predict these differences occurred simply because the participant was instructed to focus attention on the phonological structure of the stimuli and the accuracy of response. A modified version of a similar paradigm with a concurrent distraction task to shift the attention may be necessary to further confirm whether shifted attention reduces the vigilance of the monitor and yields similar uniqueness point effects in AWS. The VRT hypothesis may predict that a higher threshold for preverbal sequences that resulted in diminished uniqueness point effects for AWS was the result of anticipation of difficulty during the task by AWS participants. To further test this model, additional variables that increase the level of anticipated difficulty would be necessary, such as the inclusion of feared words as stimuli for AWS, collection of concurrent physiological data, or perhaps correlational data between communicative attitudes and experimental performance. The SAMI model may predict that the monitoring system in AWS was susceptible to state or trait factors that interfere with the typical ability of the monitor to identify critical phonemes within the preverbal code. To test this assumption, assessment or simulation of neural activity in regions associated with self-monitoring before, during, and after responses in phoneme monitoring trials may be warranted while also manipulating cognitive stress (e.g., increasingly shorter time provided for responses, performance feedback between trials, third-party or group evaluation).

A final prediction of these 3 models is that the atypical patterns observed for AWS in our study are etiological in nature rather than an adaptive pattern acquired across a lifetime of stuttering. The VRT [31], VCH ([32]; however see Howell [11] and Bernstein Ratner and Wijnen [65] for further discussion), and to a lesser extent SAMI [30] suggest that atypical monitoring in persons who stutter likely originate during childhood due to latent linguistic weaknesses or unique temperamental factors near onset. The present study was restricted to adults and not designed to address whether the differences are present in children who stutter. Instead, our data provide preliminary evidence that when adults were instructed to actively monitor inner speech – an act that may have forced participants to recruit the perceptual system more so than naturalistic speech – self-monitoring abilities were dissimilar in AWS and AWNS. To further investigate this prediction, thorough examination of uniqueness point effects in children who stutter when monitoring internal and external speech is necessary.

Perception-Based versus Production-Based Theories of Self-Monitoring

Theoretical accounts of how, and at what age, a speaker is able to self-monitor speech production have not reached consensus. Two conceptual debates should be noted during interpretation of data in the present study. First, we interpreted the differences observed when silently monitoring phonemes to reflect atypical preverbal monitoring similar to Ozdemir et al. [33]. This interpretation relies on the assumption that the preverbal monitor is governed by the speech comprehension system (e.g., [29], [35]), and this assumption has been challenged. The most frequently cited evidence against this claim comes from individuals with aphasia that exhibit either (a) poor self-monitoring but intact ability to detect other's speech errors, or (b) poor speech comprehension but intact ability to detect their own speech errors (see [66, 67] for review). If, as suggested by Levelt et al. [29], uniqueness point effects observed when monitoring inner speech occur based on the behavior of the speech comprehension system, one would also expect phoneme detection differences to be observed when monitoring phonemes in external speech. There are data available that suggest weaker auditory perception of external speech for AWS (e.g., [68]) and children who stutter (e.g., [69]). To date, only Sasisekaran and de Nil [39] have specifically examined phoneme monitoring latencies of external speech by AWS and AWNS. No significant differences when monitoring phonemes located at C_2 and C_4 positions of CVCCVC targets were observed for AWS (or AWNS). Several methodological differences between their study and the present study may have contributed to the conflicting findings. For example, it remains unclear whether uniqueness point effects would be affected by word boundaries or metrical stress patterns included within their stimuli (i.e., noun phrases: e.g., "*hot dog*" and compound words: e.g., "*hotdog*"). Nonetheless, Sasisekaran and de Nil's findings appear to contradict perception deficits in individuals who stutter (e.g., [68, 69]), as well as those of Arnstein et al. [42], and challenge the notion of perception-based monitoring deficits in AWS.

One alternative account is that the self-monitor is production-based rather than perception-based. Lu et al. [70] found significantly longer latencies for AWS during vowel judgment tasks, but not consonant judgment tasks, than AWNS. However, silent phoneme judgment latencies for consonants and vowels in AWS were strongly correlated with neural regions related to speech motor production rather than auditory processing. A production-based account of the present data is certainly plausible. For example, the phonological encoding system in AWS may have been slower to activate the remaining sublexical information even after efficient, successful isolation of the lexical item from competitors by the monitor. To reiterate, the purpose of the present study was to replicate and extend the findings of Ozdemir et al. [33] in AWS rather than clarify this conceptual debate, and behavioral data may be insufficient to attribute causality exclusively to perception or exclusively to production. Nonetheless, in light of the findings of Lu et al. [70], it is possible that the differences observed in the present study support the larger body of data that implicate speech production issues in persons who stutter [12–25], rather than the perceptual monitor itself. Further neurophysiological examination may reveal that differences in monitoring in AWS, or lack thereof, are more closely associated with difficulties in perception rather than production in AWS.

Second, it remains unclear whether the processes used to complete silent phoneme monitoring, and other metalinguistic tasks which require focused attention on inner speech, are identical to those used to monitor naturalistic speech prior to production. Huettig and Hartsuiker [71] tested this assumption in AWNS by measuring another well-documented perceptual effect (e.g., eye movement toward phonologically related stimuli) during internal monitoring tasks that did not include a metalinguistic component (e.g., visual display eye tracking). Findings were not consistent with the conclusion of Ozdemir et al. [33] that the perceptual system controls passive evaluation of inner speech. However, the authors acknowledged that the atypical, metalinguistic nature of silent phoneme monitoring tasks may have prompted speakers to engage the perceptual system. That is, speakers can still recruit the perceptual system to focus attention to inner speech, if necessary, based on the demands of the task. If this is the case, studies that require AWS and AWNS to perform metalinguistic tasks may still be confounded by atypical perceptual processes. This may be especially true for AWS who, unlike AWNS, may actively and silently "prepare" speech long before overt production in anticipation of speech difficulties. For example, Jackson et al. [72] found that all 30 AWS interviewed in their study consciously anticipate stuttered speech prior to its occurrence and, in response, often "scan" speech to avoid certain sounds (87%), or actively "focus on speech" (37%) prior to production. These well-documented preparatory behaviors may make everyday speech for AWS more metalinguistic in nature, and therefore require more active self-monitoring via the perceptual system. That is, unlike AWNS, involvement of the perceptual system when self-monitoring speech may be the norm rather than the exception for AWS.

Limitations and Future Studies

Findings from the present study are preliminary and there are, at least, a few critical limitations that should be taken into account when interpreting the outcomes and when completing future related research.

Vocabulary

Participants with larger vocabularies may require more time to eliminate a greater number of competitors, which may in turn influence their response to the uniqueness point of specific stimuli. The impact of intraparticipant variability in vocabulary knowledge relative to the specific target words would be difficult to capture using formal vocabulary assessments due to (a) unlikely inclusion of items phonologically similar to specific targets, and (b) the exclusion of phonologically similar proper nouns. In addition, although all participants in the present study reported native English skills and limited knowledge of additional languages per the LHQ [54], 10 reported completion of foreign language coursework during their academic career. It should be noted that of these 10 participants, all reported that current use of any foreign language was minimal (i.e., less than 15 min or 0 min per month, on average, in any modality). Further, inclusion of proficiency in any language other than English as an additional covariate did not affect the critical, significant interaction. Nevertheless, we cannot completely assume that non-English targets with similar phonological stems were also coactivated during our experimental task. One approach to address these concerns in future studies would be postexperimental assessment wherein participants offer any additional words they can recall, in any language, based on the phonological stem of each target word.

Gender

Ideally, a perfect balance of males and females within each group would minimize any potential gender-based bias. In the present study, AWNS were predominately female and AWS, as expected, were predominately male [73]. Although there are no data to suggest the speed of monitoring phonemes during inner speech would differ based on gender, future studies should strive to obtain larger sample sizes with similar gender ratios to eliminate, or elucidate, any potential response bias between groups.

Restricted Stimuli

Obviously, a greater number of target items per condition would have been preferable to increase statistical power. Ozdemir et al. [33] noted this limitation as well when discussing the difficulty of finding adequate tokens in Dutch (e.g., 10 EUP, 10 LUP, and 10 NUP). Given the necessary constraints of stimuli selection, it was not possible to find additional lexical items in Standard American English. Additional LUP and EUP stimuli provided in previous studies by Gaskell and Marslen-Wilson [47] and Henderson et al. [64] were not eligible for the silent phoneme monitoring task described in the present study for several reasons. First, we limited stimuli to bisyllabic, 6-segment targets with trochaic stress to ensure comparable phonological distance from onset to word-final phonemes without additional metrical processing. Second, stimuli in silent phoneme monitoring tasks, unlike cross-modal priming or pause-detection tasks, cannot contain repeated consonants as targets. Third, participants in the present study were required to silently generate words via picture. Many of the words used in their studies would not be imageable (e.g., *device*, *passage*, *surface*) and therefore ineligible as visual stimuli. Fourth and finally, inclusion of EUP and LUP tokens standardized in British English dialect may have influenced the speed or accuracy of response of Standard American English participants, particularly for LUP stimuli wherein the uniqueness point phoneme was the unstressed medial vowel (see Table 2). Although these limitations restricted the breath of our targets, we feel that stimuli selected for the present study had certain advantages relative to previous studies. For example, lexical competitors were lower in frequency than Henderson et al. [64] (e.g., target: *camera*, competitors: *camouflage*, *camel*; target: *salad*, competitors: *salary*, *salamander*), and therefore less likely to generate lexical competition and strengthen the desired uniqueness point effect, if present (see Appendix). Nonetheless, to address this potential limitation, future studies should assess uniqueness point effects in AWS using auditory targets rather than silently generated targets to ease these restrictions, namely imageability, and increase the number of eligible tokens.

Conclusion

Preliminary data from the present study indicate that perception-specific patterns observed in external and preverbal speech monitoring in AWNS – specifically the uniqueness point effect – were not observed for AWS. The absence of these expected patterns supports theories of stuttering that implicate deficits in the preverbal speech monitor. Future studies wherein participants monitor internal *and* external speech, and at younger ages, are necessary to confirm these preliminary findings and lend further support to monitor deficit theories of stuttering.

Acknowledgements

We would like to thank the Michael and Tami Lang Stuttering Institute and Dr. Jennifer and Emanuel Bodner for their endowed support of our research efforts. This project was supported by the LSU-ORED Faculty Research Grant and Manship Summer Stipend Award, as well as the Louisiana Board of Regents Research Competiveness Subprogram (LEQSF[2015–18]-RD-A-03). Most of all, we would like to thank the adults who do and do not stutter willing to participate.

Disclosure Statement

The authors have no conflict of interest to declare.

Appendix

Lexical Competitors and Corresponding Word Frequency for Target Stimuli at Specified Uniqueness Point

	IPA	UP	WF	Lexical competitors (WF)
Early uniqueness point				
Dolphin	/d ɑ l f ɪ n/	/f/	2.76	dolphinarium (N/A); dolphin-safe (N/A), Dolphus (N/A)
Falcon	/f æ l k ə n/	/k/	3.31	falcate (N/A); Falonbridge (N/A); Falconcrest (N/A); falconer (N/A); falconry (N/A), Falkner (N/A)
Napkin	/n æ p k ɪ n/	/k/	3.61	none
Tadpole	/t æ d p o l/	/p/	0.59	none
Walrus	/w ɑ l r ə s/	/r/	1.12	Waller (0.51)
Late uniqueness point				
Biscuit	/b ɪ s k ɪ t/	/ɪ/	3.75	none
Closet	/k l ɑ z ə t/	/ə/	27.08	none
Forest	/f o r ə s t/	/s/	18.88	forester (N/A); Forrester (N/A)
Spinach	/s p ɪ n ɪ ʧ/	/ɪ/	2.55	spinifex (N/A)
Stomach	/s t ^ m ə k/	/ə/	33.82	stomachache (1.06); stomach-churning (N/A); stomacher (N/A); stomachful (N/A); stomachic (N/A)
No uniqueness point				
Bandage	/b æ n d ɪ ʤ/	/ʤ/	2.86	none
Compass	/k ^ m p ə s/	/s/	4.06	none
Magnet	/m æ g n ə t/	/t/	2.75	magnetism (0.96); magnetite (N/A); magnetization (N/A); magnetize (N/A); magnetron (N/A)
Mattress	/m æ t r ə s/	/s/	6.61	none
Mustard	/m ^ s t ɚ d/	/d/	6.45	none

IPA, transcription in International Phonetic Alphabet; UP, uniqueness point phoneme identified using the Sound Search function in the Longman Pronunciation Dictionary – Third Edition [57]. WF, Word frequency values per million words calculated using CLEARPOND lexical database [58]. Lexical competitors also identified using Longman [57] database. N/A, word frequency not available in CLEARPOND database.

References

1 Mersov A-M, Jobst C, Chyne DO, de Nil L: Sensorimotor oscillations prior to speech onset reflect altered networks in adults who stutter. Front Hum Neurosci 2016;10:443.

2 Namasivayam AK, van Lieshout P: Speech motor skills and stuttering. J Mot Behav 2011; 43:477–489.

3 Howell P: The EXPLAN theory of fluency control applied to the treatment of stuttering; in Fava E (ed): Clinical Linguistics: Theory and Applications in Speech Pathology and Therapy. Philadelphia, J Benjamins Publishing, 2002, pp 95–118.

4 Postma A, Kolk H: The covert repair hypothesis: prearticulatory repair processes in normal and stuttered disfluencies. J Speech Lang Hear Res 1993;36:472–487.

5 Smith A, Kelly E: Stuttering: A dynamic, multifactorial model; in Curlee RF, Siegel GM (eds): Nature and Treatment of Stuttering: New Directions. Needham Heights, Allyn and Bacon, 1997, pp 204–217.

6 Walden TA, Frankel CB, Buhr AP, Johnson KN, Conture EG, Karrass JM: Dual diathesis-stressor model of emotional and linguistic contributions to developmental stuttering. J Abnorm Child Psychol 2012;40:663–644.

7 Arndt J, Healey EC: Concomitant disorders in school-age children who stutter. Lang Speech Hear Serv Sch 2001;32:68–78.

8 Blood G, Ridenour VJ, Qualls CD, Hammer CS: Co-occurring disorders in children who stutter. J Commun Disord 2003;36:428–448.

9 Pelczarski KM, Yaruss JS: Phonological encoding of young children who stutter. J Fluency Disord 2014;39:12–24.

10 Nippold MA: Stuttering and phonology: is there an interaction? Am J Speech Lang Pathol 2002;11:99–110.

11 Howell P: Recovery from Stuttering. New York, Taylor and Francis, 2011.

12 Ambrose NG, Yairi E, Loucks TM, Seery CH, Throneburg R: Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: initial findings. J Fluency Disord 2015;45: 12–26.

13 Paden EP, Yairi E: Phonological characteristics of children who stuttering persisted or recovered. J Speech Lang Hear Res 1996;39: 981–990.

14 Spencer C, Weber-Fox C: Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence from stuttering. J Fluency Disord 2014; 41:32–46.

15 Mohan R, Weber C: Neural systems mediating processing of sound units of language distinguish recovery versus persistence in stuttering. J Neurodev Disord 2015;7:28.

16 Byrd CT, Vallely M, Anderson J, Sussman H: Nonword repetition and phoneme elision in adults who do and do not stutter. J Fluency Disord 2012;37:188–201.

17 Byrd CT, McGill M, Usler E: Nonword repetition and phoneme elision in adults who stutter: vocal and nonvocal performance differences. J Fluency Disord 2015;44:17–31.

18 Coalson GA, Byrd CT: Nonword repetition in adults who stutter: the effects of stimuli stress and auditory-orthographic cues. PLoS One 2017;12:e0188111.

19 Brocklehurst PH, Corley M: Investigating the inner speech of people who stutter: evidence for (and against) the Covert Repair Hypothesis. J Commun Disord 2011;44:246–260.

20 McGill M, Sussman H, Byrd CT: From grapheme to phonological output: performance of adults who stutter on a word jumble task. PLoS One 2016;11:e0151107.

21 Jones RM, Fox RA, Jacewicz E: The effects of concurrent cognitive processing load on phonological processing in adults who stutter. J Speech Lang Hear Res 2012;55:1862–1875.

22 Weber-Fox C, Spencer RM, Spruill JE, Smith A: Phonological processing in adults who stutter: electrophysiological and behavioral evidence. J Speech Lang Hear Res 2004;47: 1244–1258.

23 Coalson GA, Byrd CT: Metrical encoding in adults who do and do not stutter. J Speech Lang Hear Res 2015;58:601–621.

24 Sasisekaran J, de Nil LF, Smyth R, Johnson C: Phonological encoding in the silent speech of persons who stutter. J Fluency Disord 2006; 31:1–21.

25 Sasisekaran J, Brady A, Stein J: A preliminary investigation of phonological encoding in children who stutter. J Fluency Disord 2013; 38:45–58.

26 Burger R, Wijnen F: Phonological encoding and word stress in stutter and nonstuttering adults. J Fluency Disord 1999;24:91–106.

27 Vincent I, Grela BG, Gilbert HR: Phonological priming in adults who stutter. J Fluency Disord 2012;37:91–105.

28 Hennessey NW, Nang C, Beilby JM: Speeded verbal responding in adults who stutter: are there deficits in linguistic encoding? J Fluency Disord 2008;33:180–202.

29 Levelt WJM, Roelofs A, Meyer AS: A theory of lexical access in speech production. Behav Brain Sci 1999;22:1–75.

30 Arenas RM: Conceptualizing and investigating the contextual variability of stuttering: the speech monitoring interaction (SAMI) framework. Speech Lang Hear 2017;20:15–28.

31 Brocklehurst PH, Lickley RJ, Corley M: Revisiting Bloodstein's Anticipatory Struggle Hypothesis from a psycholinguistic perspective: a Variable Release Threshold hypothesis of stuttering. J Commun Disord 2013;46:217–237.

32 Vasic N, Wijnen F: Stuttering as a monitoring deficit; in Hartsuiker RJ, Bastiaanse R, Postma A, Wijnen F (eds): Phonological Encoding and Monitoring in Normal and Pathological Speech. New York, Psychology Press, 2005, pp 226–247.

33 Ozdemir R, Roelofs A, Levelt WJM: Perceptual uniqueness point effects in monitoring internal speech. Cognition 2006;105:457–465.

34 Indefrey P, Levelt WJM: The spatial and temporal signatures of word production components. Cognition 2004;92:101–144.

35 Hartsuiker RJ: Studies on verbal self-monitoring: the perceptual loop model and beyond; in Meyer AS, Wheeldon LR, Krott A (eds): Automaticity and Control in Language Processing. Brussels, Psychology Press, 2007, pp 93–121.

36 Wheeldon LR, Levelt WJM: Monitoring the time course of phonological encoding. J Mem Lag 1995;34:311–334.

37 Wheeldon LR, Morgan JL: Phoneme monitoring in internal and external speech. Lang Cogn Process 2002;17:503–535.

38 Connine CM, Titone D: Phoneme monitoring. Lang Cogn Process 1996;11:635–645.

39 Sasisekaran J, de Nil LF: Phoneme monitoring in silent naming and perception in adults who stutter. J Fluency Disord 2006;31:284–302.

40 Sasisekaran J: Exploring the link between stuttering and phonology: a review and implications for treatment. Semin Speech Lang 2014;35:95–113.

41 Lickley RJ, Hartsuiker RJ, Corley M, Russell M, Nelson R: Judgment of disfluency in people who stutter and people who do not stutter: result from magnitude estimation. Lang Speech 2005;48:299–312.

42 Arnstein D, Lakey B, Compton RJ, Kleinow J: Preverbal error-monitoring in stutterers and fluent speakers. Brain Lang 2011;116:105–115.

43 Postma A, Kolk H: Error monitoring in people who stutter: evidence against auditory feedback defect theories. J Speech Lang Hear Res 1992;35:1024–1032.

44 Postma A, Kolk H, Povel DJ: Speech planning and execution in stutterers. J Fluency Disord 1990;15:49–59.

45 Marslen-Wilson W: Function and process in spoken word recognition: A tutorial review; in Bouma H, Bouwhuis DG (eds): Attention and Performance X: Control of Language Processes. Hillsdale, Erlbaum, 1984.

46 Marslen-Wilson W: Activation, competition, and frequency in lexical access; in Altmann G (ed): Cognitive Models of Speech Processing. Cambridge, MIT Press, 1990, pp 148–172.

47 Gaskell MG, Marslen-Wilson WD: Representation and competition in the perception of spoken words. Cogn Psychol 2002;45:220–266.

48 Frauenfelder UH, Segui J, Dijkstra T: Lexical effects in phonemic processing: facilitatory or inhibitory? J Exp Psychol Hum Percept Perform 1990;16:77–91.

49 Pitt MA, Samuel AG: Lexical and sublexical feedback in auditory word recognition. Cogn Psychol 1995;29:149–188.

50 Wagner RK, Torgesen JK, Rashotte CA: Comprehensive Test of Phonological Processes (CTOPP). Austin, Pro-Ed, 1999.
51 Wagner RK, Torgesen JK, Rashotte CA, Pearson NA: Comprehensive Test of Phonological Processes – Second Edition (CTOPP-2). Austin, Pro-Ed, 2013.
52 American Speech-Language-Hearing Association: Guidelines for audiologic screening (Guidelines). 1997. www.asha.org/policy.
53 Oldfield RC: The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia 1971;9:97–113.
54 Li P, Sepanski S, Zhao X: Language history questionnaire (LHQ): a Web-based interface for bilingual research. Behav Res Methods 2006;38:202–210.
55 Riley GD: Stuttering Severity Instrument – 4th Edition (SSI-4). Austin, Pro-Ed, 2009.
56 Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, Neely JH, Nelson GB, Treiman R: The English Lexicon Project. Behav Res Methods 2007;39:445–459.
57 Wells JC: Longman Pronunciation Dictionary, ed 3. Essex, Pearson Education, 2008.
58 Marian V, Bartolotti J, Chabal S, Shook A: CLEARPOND: cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. PLoS One 2012; 7:e43230.
59 Brysbaert M, New B: Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measures for American English. Behav Res Methods 2009;41:977–990.
60 Coltheart M: The MRC psycholinguistic database. Q J Exp Psychol 1981;33:497–505.
61 Szekely A, Jacobsen T, D'Amico S, Devescovi A, Andonova E, Herron D, Wicha N: A new on-line resource for psycholinguistic studies. J Mem Lang 2004;51:247–250.
62 Yairi E, Ambrose N: Early Childhood Stuttering: For Clinicians, by Clinicians. Austin, Pro-Ed, 2005.
63 Natke U, Sandrieser P, van Ark M, Pietrowsky R, Kalvernam KT: Linguistic stress, within-word position, and grammatical class in relation to early childhood stuttering. J Fluency Disord 2004;29:109–122.
64 Henderson L, Weighall A, Brown H, Gaskell G: Online lexical competition during spoken word recognition and word learning in children and adults. Child Dev 2013;84:1668–1685.
65 Bernstein Ratner N, Wijnen F: The vicious cycle: Linguistic encoding, self-monitoring and stuttering; in Au-Yeung J (ed): Proceedings of the 5th World Congress of Fluency Disorders. Dublin, International Fluency Association, 2007, pp 84–90.
66 Hartsuiker RJ: Monitoring and control of the production system; in Goldrick M, Ferreira V, Miozzo M (eds): The Oxford Handbook of Language Production. New York, Oxford University Press, 2014, pp 417–436.
67 Nozari N, Dell GS, Schwartz MF: Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. Cogn Psychol 2011;63:1–33.
68 Neef NE, Sommer M, Neef A, Paulus W, von Gudenberg AW, Jung K, Wustenberg T: Reduced speech perceptual acuity for stop consonants in individuals who stutter. J Speech Lang Hearing Res 2012;55:276–289.
69 Jansson-Verkasalo E, Eggers K, Jarvenpaa A, Suominen K, van den Bergh B, de Nil L, Kujala T: Atypical central auditory speech-sound discrimination in children who stutter as indexed by the mismatch negativity. J Fluency Disord 2014;41:1–11.
70 Lu C, Long Y, Zheng L, Shi G, Ding, G, Howell P: Relationship between speech production and perception in people who stutter. Front Hum Neurosci 2016;10:224.
71 Huettig F, Hartsuiker RJ: Listening to yourself is like listening to others: external, but not internal, verbal self-monitoring is based on speech perception. Lang Cogn 2010;25:347–374.
72 Jackson ES, Yaruss JS, Quesal RW, Terranova V, Whalen DH: Responses of adults who stutter to the anticipation of stuttering. J Fluency Disord 2015;45:38–51.
73 Bloodstein O: A Handbook on Stuttering. San Diego, Singular, 1995.