Phonological working memory in developmental stuttering: Potential Insights from the neurobiology of language and cognition

4 authors, including:

Andrew Bowers
University of Arkansas
16 PUBLICATIONS 173 CITATIONS

Lisa M. Bowers
University of Arkansas
16 PUBLICATIONS 95 CITATIONS

Daniel Hudock
Idaho State University
16 PUBLICATIONS 59 CITATIONS
Phonological working memory in developmental stuttering: Potential insights from the neurobiology of language and cognition

Andrew Bowers, Lisa M. Bowers, Daniel Hudock, Heather L. Ramsdell-Hudock

University of Arkansas, Epley Center for Health Professions, 606 N. Razorback Road, Fayetteville, AR 72701, United States
Idaho State University, 650 Memorial Dr. Bldg. 68, Pocatello, ID 83201, United States

ARTICLE INFO

Keywords:
Phonological working memory
Sensorimotor integration
Executive function
Nonword repetition
Dual-task
Stuttering

ABSTRACT

The current review examines how neurobiological models of language and cognition could shed light on the role of phonological working memory (PWM) in developmental stuttering (DS). Toward that aim, we review Baddeley’s influential multicomponent model of PWM and evidence for load-dependent differences between children and adults who stutter and typically fluent speakers in nonword repetition and dual-task paradigms. We suggest that, while nonword repetition and dual-task findings implicate processes related to PWM, it is unclear from behavioral studies alone what mechanisms are involved. To address how PWM could be related to speech output in DS, a third section reviews neurobiological models of language proposing that PWM is an emergent property of cyclic sensory and motor buffers in the dorsal stream critical for speech production. We propose that anomalous sensorimotor timing could potentially interrupt both fluent speech in DS and the emergent properties of PWM. To further address the role of attention and executive function in PWM and DS, we also review neurobiological models proposing that prefrontal cortex (PFC) and basal ganglia (BG) function to facilitate working memory under distracting conditions and neuroimaging evidence implicating the PFC and BG in stuttering. Finally, we argue that cognitive-behavioral differences in nonword repetition and dual-tasks are consistent with the involvement of neurocognitive networks related to executive function and sensorimotor integration in PWM. We suggest progress in understanding the relationship between stuttering and PWM may be accomplished using high-temporal resolution electromagnetic experimental approaches.

Abbreviations: DS, developmental stuttering; CWS, children who stutter; AWS, adults who stutter; PWM, phonological working memory; PFC, prefrontal cortex; BG, basal ganglia; LTM, long term memory; PL, phonological loop; VS, visuospatial sketchpad; CE, central executive; EB, episodic buffer; TFS, typically fluent speakers; PWS, people who stutter; CWNS, children who do not stutter; ERP, event-related potential; BGTC, basal ganglia thalamo cortical; IFG, inferior frontal gyrus; PMC, premotor cortex; PT, planum temporale; Spt, sylvian parieto-temporal junction; STG, superior temporal gyrus; STS, superior temporal sulcus; MTG, middle temporal gyrus; SFC, state feedback control; CA, conduction aphasia; BA, Brodman’s Area; IPL, inferior parietal lobule; SPL, superior parietal lobule; PBWM, prefrontal cortex basal ganglia working memory; SMA, supplemental motor area; ACC, anterior cingulate cortex; DIVA, directions in auditory space to velocities in articulator space; GODIVA, Gradient order DIVA model; vMC, ventral motor cortex; fMRI, functional magnetic resonance imaging; SMG, supramarginal gyrus; DMN, default mode network; DAN, dorsal attention network; VAN, ventral attention network; FPN, fronto-parietal network; GMV, gray matter volume; MFG, medial frontal gyrus; FA, fractional anisotropy; PET, positron emission tomography; ALE, activation likelihood estimation; A1, auditory cortex; OFC, orbitofrontal cortex; SSI-4, Stuttering Severity Instrument – 4th Edition; EEG, electroencephalography; MEG, magnetoencephalography; SMS, Speech Motor Skill

E-mail addresses: albowers@uark.edu (A. Bowers), lmbowers@uark.edu (L.M. Bowers), hudock@isu.edu (D. Hudock), ramsdell@isu.edu (H.L. Ramsdell-Hudock).

https://doi.org/10.1016/j.jfludis.2018.08.006

Received 29 September 2017; Received in revised form 30 July 2018; Accepted 27 August 2018
Available online 01 September 2018
0094-730X/ Published by Elsevier Inc.
1. Introduction

Developmental stuttering (DS) is a disorder of speech fluency associated with a high recovery rate in preschool-age children who stutter (CWS) (> 80%) and significant psychosocial and emotional consequences for those who persist into later childhood and adulthood. Multifactorial, dynamic models propose that DS is primarily a disorder affecting the sensorimotor control of speech that is influenced by other psycholinguistic and cognitive-emotional factors (Smith, 1999; Smith & Weber, 2017; Starkweather, 2002; van Lieshout, Hulstijn, & Peters, 2004). From this point of view, interacting general cognitive capacities such as working memory, attention, and executive function play some role in the moment-to-moment variability and phonotypic expression of stuttering in both CWS (Chang et al., 2017; Pelczarski & Yaruss, 2016; Smith & Weber, 2017; Spencer & Weber-Fox, 2014) and adults who stutter (AWS) (Alm, 2004, Alm & Risberg, 2007; Kell et al., 2017; van Leishout et al., 2004). In particular, cognitive capacities known to support phonological working memory (PWM), including phonological storage and the articulatory rehearsal of speech sounds, have been implicated as a factor in preschool CWS that predicts subsequent recovery or persistence (Mohan & Weber, 2015; Spencer & Weber-Fox, 2014). In CWS, PWM capacity may also be related to disturbances in cognitive-linguistic processing and to variability in the expression of stuttering in contexts that load other cognitive processes supporting it, including attentional control and executive function both in CWS (Bajaj, 2007; Eggers et al., 2012; Jones et al., 2017; Usler & Weber-Fox, 2015) and AWS (Alm, 2004; Alm & Risberg, 2007; Bajaj, 2007; Bosshardt, 2006). As such, a better understanding of the cognitive processes underlying PWM has the potential to address long held questions about the nature of stuttering.

Despite the potential significance of PWM, studies using tasks thought to load PWM in CWS and AWS have produced somewhat mixed results, and the underlying mechanisms mediating recently reported differences remain unclear (Anderson & Wagovich, 2010; Anderson et al., 2006; Byrd et al., 2015; Byrd et al., 2012; Smith et al., 2010). Besides as a consequence, while both CWS and AWS present with functional and structural anomalies in both sensorimotor (Buchsbaum & D’Esposito, 2008; Buchsbaum et al., 2011; Hickok et al., 2011; Jacquemot & Scott, 2006) and neurocognitive networks implicated in PWM (Chang et al., 2017; Kell et al., 2017; Sitek et al., 2016), there have been no attempts to integrate existing cognitive-behavioral and neuroimaging findings with neurobiological models of language and cognition. As such, the aim of the current review is to examine how neurobiological models incorporating PWM in typical speech and language processing (Hickok & Poeppel, 2007; Jacquemot & Scott, 2006) and executive function (D’Esposito, 2007; Hazy et al., 2007) may shed light on underlying mechanisms mediating recent cognitive-behavioral studies implicating lower PWM capacity in CWS (Anderson & Wagovich, 2010; Anderson et al., 2006; Spencer & Weber-Fox, 2014) and AWS (Byrd et al., 2015; Byrd et al., 2012; Smith et al., 2010). We argue that interpreting cognitive-behavioral findings in the context of neurobiological models may strengthen the relationship between underlying mechanisms affecting fluent speech production those underlying differences in PWM in both CWS (Spencer & Weber-Fox, 2014) and AWS (Byrd et al., 2015; Byrd et al., 2012).

Toward bridging the gap between cognitive-behavioral findings in DS and the neurobiology of PWM, Section 2 reviews the relationship between DS and PWM capacity in the context of Baddeley’s influential multicomponent working memory model and behavioral findings from nonword repetition and dual-task paradigms that implicate aspects of the executive function and phonological storage in CWS and AWS. In Section 3, we review neurobiological models proposing that PWM is an emergent property of cyclic input and output buffers in the dorsal sensorimotor stream (Buchsbaum & D’Esposito, 2008; Hickok et al., 2011; Jacquemot & Scott, 2006) and models proposing that interactions between the prefrontal cortex (PFC) and basal ganglia (BG) mediate working memory in attentionally demanding conditions (D’Esposito, 2007; Hazy et al., 2007). In support of differences in networks supporting PWM in DS, we review neuroimaging evidence implicating large-scale networks in both CWS and AWS, including PFC, BG, and dorsal stream sensorimotor regions. In Section 4, we posit that mistimed or ‘noisy’ dorsal stream sensorimotor integration previously proposed to cause stutter-typical disfluencies (i.e., part-word repetitions, prolongations, and blocks) may also account for load-dependent differences on nonword repetition tasks in CWS and AWS. We further suggest that increased executive control of speech may play a modulatory role in regulating dorsal stream timing, as experience with stuttering accumulates over the course of speech and language development. Lastly, in Section 5, we propose that progress in understanding the cognitive processes underlying PWM in DS may be accomplished using noninvasive, high temporal resolution time-frequency methods in nonword repetition and dual-task paradigms.

2. Phonological working memory in DS

Much of the most recent work investigating aspects of PWM in DS references Baddeley’s working memory model as an explanatory framework and thus a brief review of the model is worthwhile (Anderson & Wagovich, 2010; Anderson et al., 2006; Bajaj, 2007; Bosshardt, 2006; Byrd et al., 2015; Spencer & Weber-Fox, 2014). Currently, the model consists of four primary components proposed to interact with long-term memory (LTM), which include the phonological loop (PL), visuospatial sketchpad (VS), central executive (CE), and episodic buffer (EB) (Baddeley, 2003, 2012). The PL is arguably the component most relevant to the acquisition of speech and language and is composed of two subcomponents, one component that provides temporary storage for phonological input known as the phonological store and a second to refresh the first via subvocal articulatory rehearsal. Functionally, the PL is proposed to support new phonological learning, facilitate the acquisition of new vocabulary, support second-language acquisition, and potentially influence action selection via subvocal rehearsal (Adams & Gathercole, 1995; Alloway et al., 2005; Baddeley, 2012). The VS functions to integrate spatial, visual, and kinesthetic information into a single representation for temporary storage. The CE component provides access to the other components and is broadly tasked with the control of attention to items to be held in working memory. The latest and least studied addition to the model is the EB, a process proposed to have the capacity to store
2.1. Review of the evidence: nonword repetition in DS

A growing corpus of experimental and descriptive evidence implicates PWM differences in both CWS and AWS as compared to typically fluent speakers (TFS). Much of the evidence linking PWM with stuttering has come from what are known as nonword repetition paradigms. Nonword repetition involves repeating words that conform to phonological rules in the language but have no meaning (i.e., semantic content). In Baddeley’s model, such a task is thought to tap the storage component of the PL without reliance on LTM for language (Gathercole et al., 1994). Following that reasoning, differences between people who stutter (PWS) and TFS on nonword repetition tasks in the absence of differences on verbal repetition tasks would provide evidence favoring differences in PWM as distinct from LTM for semantic content. Findings from nonword repetition tasks are of importance because nonword repetition capacity in preschool children, as distinct from verbal working memory tasks, has been reported to predict subsequent recovery or persistence (Spencer & Weber-Fox, 2014). As such, in the sections that follow, we first review studies reporting both significant group differences and a lack of differences in nonword repetition accuracy and subsequently discuss factors that may have led to nonsignificant findings in some studies. We suggest that the weight of the evidence implicates subtle, load-dependent differences in PWS from the preschool years up through adulthood.

2.1.1. Nonword repetition findings in preschool-age CWS

Early studies investigating nonword repetition in CWS provided somewhat mixed results with two studies in CWS (age range 2–8 years) reporting significantly lower behavioral performance on children’s test of nonword repetition (Gathercole et al., 1994), a common measure of nonword repetition (Anderson et al., 2006; Hakim & Ratner, 2004). In a follow-up study, Anderson and Wagovich (2010) reported that preschool CWS (3–5 years) who participated in both a picture naming task and a nonword repetition task were significantly lower on the nonword repetition task only compared to children who do not stutter (CWNS). Interestingly, the two groups did not differ on a common measure of attention via parental report. However, within the group of CWS a measure of ‘attentional focusing’ was positively correlated with nonword repetition performance. More recently, Smith, Goffman, Sasisekaran, and Weber-Fox (2012) did not detect group differences on a nonword repetition task in a group of preschool-age CWS (4–6 years). In that study, differences were only apparent for CWS with concomitant speech-sound and/or language disorders, suggesting that other articulatory and cognitive-linguistic capacities may interact to influence performance on nonword repetition tasks. Further, although CWS without speech-sound/language disorders did not differ on a measure of nonword repetition accuracy, they did show higher variability on indices of oral motor coordination (i.e., lip aperture variability), suggesting that motor control of speech varied with nonword repetition load.

In contrast to Smith et al. (2012), Spencer and Weber-Fox (2014) did find significant differences in nonword repetition performance between a preschool group of CWS and CWNS (age range 3–5 years) with or without concomitant language and phonological disorders. Additionally, findings revealed that nonword repetition performance was significantly lower for CWS who persisted relative to those who recovered. Importantly, neither measures of expressive/receptive language nor measures of working memory for words with semantic content were associated with recovery or persistence, suggesting a disassociation between verbal working memory for words with semantic content and PWM. Most recently, Pelczarski and Yaruss (2016) reported statistically significant yet subclinical differences on a norm-referenced test of nonword repetition in CWS (age range 5–6 years). Consistent with previous studies, no differences were reported on a digit-span task, suggesting similar behavioral performance for words with semantic content. Taken together, the majority of nonword repetition studies in preschool CWS (5/6) suggest subtle, subclinical differences that may also be related to or interact with articulation, phonology, and attentional capacities. Further, studies that have also investigated aspects of verbal working memory via digit span or other memory tasks for words with semantic content also suggest that reduced capacities for nonword sequences are not related verbal working memory capacity (Pelczarski & Yaruss, 2016) and do not appear to predict recovery or persistence (Spencer & Weber-Fox, 2014). Relevant details of the studies cited above, including sample sizes, maximum syllable length of nonword stimuli, and effect sizes when reported are summarized in Appendix A (Table A1).

2.1.2. Nonword repetition findings in school-age CWS

Relative to findings in preschool CWS, fewer published studies have investigated differences in school-age CWS and CWNS (age range 5–15 years). However, unlike studies of preschool CWS, only 3/6 studies report significantly lower performance in the school-age CWS group when compared to CWNS. In the earliest report of nonword repetition in school-age CWS, Seery, Watkins, Ambrose, and Throneburg (2006) reported no significant main effect between CWS and CWNS but did report a significant interaction between group and condition (i.e., nonword syllable length) at 5 syllable nonword length. Bakhtiar, Abadm, and Panahi (2007) found no significant differences in nonword repetition accuracy between 12 CWS and 12 CWNS (age range 5–7 years) using nonwords with a maximum length of 3 syllables. A retrospective study of 5 persistent CWS and 14 recovered CWS also reported no statistically significant group differences on a nonword repetition task (Chon & Ambrose, 2007). Later, Weber-Fox, Spruill, Spencer, and Smith (2008) reported that nonword repetition accuracy in a group of 10 CWS was similar to a group of 30 CWNS in the same age range. In contrast to previous studies, Oyoun, El Dessouky, Sahar, and Fawzy (2010) did find significant differences in nonword repetition accuracy between 30 CWS and 30 matched CWNS, suggesting that previous studies using smaller sample sizes may not have detected a difference due to a lack of statistical power. A more recent study conducted by Sasisekaran and Byrd (2013) in CWS and CWNS ages...
7–15, found no significant difference between CWS and CWNS in overall nonword repetition accuracy and further found no differences in a nonword repetition condition requiring the omission of a target phoneme (i.e., phoneme elision). However, further subdivision of the CWS and CWNS into younger (7–11) and older (11–15) groups revealed a difference between the younger and older CWS that was not present in the CWNS, suggesting that age may be a particularly important factor in determining whether or not between group differences are detectable in school-age children. In summary, while most studies of school-age CWS have not reported between subject differences, those studies also used nonword stimuli of shorter syllable length than studies of preschool CWS. Further, the two studies that did detect overall lower performance in CWS, either employed a 5 syllable length condition (Seery et al., 2006) or employed a larger sample size (Oyoun et al., 2010). As such, while it is as yet unclear whether school-age CWS present with robust differences in nonword repetition accuracy, nonword repetition load, age, and sample size appear to be important factors in determining whether a difference is detectable. A table summarizing findings from the studies cited above is presented in Appendix A (Table A2).

2.1.3. Nonword repetition findings in AWS

Although most studies of school-age children have not reported significant differences in nonword repetition accuracy, most (5/7) studies in AWS have reported lower behavioral performance. An early preliminary study reported significant differences in nonword repetition performance that was consistent with significantly lower capacity in AWS relative to TFS (Ludlow et al., 1997). In a kinematic study of nonword repetition, Namasiyavam and van Lieshout (2008) reported that AWS showed reduced practice effects reflected in higher movement variability over several days when compared to controls. In another study of AWS who were categorized with mild stuttering severity, Smith et al. (2010) found no differences in accuracy on nonword repetition but did report group differences on a measure of inter-articulator coordination (i.e., labial aperture variability) with increasing length (1–4 syllables) and phonological complexity of nonwords. While earlier studies had used shorter nonword lengths, Byrd et al. (2012) investigated group differences at higher nonword loads (i.e., 7 syllables). A significant group by syllable interaction was reported at the 7 syllable length nonwords only. In a subsequent study using a group of 9 AWS and matched controls, Sasisekaran (2013) found no behavioral differences on the Nonword Repetition Test (Dollaghan & Campbell, 1998) for 1–4 syllables but did find significant group differences using a measure of lip aperture variability on a nonword reading task in which word lengths were varied between 6 and 11 syllables. Subsequently, Sasisekaran and Weisberg (2014) reported that AWS showed an increasing likelihood of nonword repetition errors with increasing length and phonological complexity of nonwords (i.e., 6 syllables), suggesting that the load induced by nonword stimuli may be a critical factor in detecting between group differences in AWS.

Building on findings from Byrd et al. (2012), Byrd et al. (2015) reported that nonword repetition performance differed between AWS and TFS on 7 syllable nonwords but no differences were reported on 4 syllable nonwords. Additionally, decreases in performance were observed on a phoneme elision condition for 4 and 7 syllable nonwords in which participants silently identified the nonword with a target phoneme omitted, strongly implicating phonological memory traces in the observed differences. Finally, a recent report also demonstrated that AWS were less accurate at repeating low-frequency iambic stress patterns in the absence of auditory-orthographic cues, suggesting that even at shorter syllable lengths phonological load manipulations may allow for the detection of group differences in behavioral accuracy (Coalson & Byrd, 2017). Thus, collectively findings manipulating nonword repetition load suggest that AWS present with lower performance on nonword repetition tasks when PWM is heavily loaded even into adulthood. A table summarizing findings from the studies cited above are presented in Appendix A (Table A3).

2.1.4. Implications of nonword repetition findings in CWS and AWS

In summary, the weight of the evidence implicates nonword repetition differences from the onset of stuttering in the preschool years up through adulthood. First, while the majority of studies in preschool CWS have detected a group difference, some studies that did not find significance may not have controlled for a number of influential factors. In the only study of preschool CWS reporting no overall group difference in accuracy, subgroups with language disorders and phonological disorders did present with significantly lower behavioral performance (Smith et al., 2012). By contrast, two other subsequent studies found group differences in preschool CWS even without clinically significant language or phonological disorders (Pelczarski & Yaruss, 2016; Spencer & Weber-Fox, 2014). As noted by Pelczarski and Yaruss (2016), age, language ability, and nonword stimulus factors such as length and complexity may not have been adequately examined in previous studies. Further, most of the studies have included preschool children that may subsequently recover, a factor that is likely to lead to higher overall behavioral performance in CWS groups (Spencer & Weber-Fox, 2014). Given the weight of the evidence and factors that may have affected findings in previous studies reporting no group differences, nonword repetition capacity is strongly implicated in preschool-age CWS (Table A4).

In school-age children who stutter, while only 3 of 6 studies detected group differences, the studies that did employed larger sample sizes (Oyoun et al., 2010), separated CWS into younger and older groups (Sasisekaran & Byrd, 2013), or detected differences only at longer nonword lengths (Seery et al., 2006). These findings suggest that school-age CWS may present with higher PWM capacities than preschool CWS and show differences primarily under high nonword repetition loads. That interpretation is consistent with available evidence in AWS suggesting that the likelihood of detecting differences in behavioral performance is heavily dependent on factors affecting nonword repetition load (Byrd et al., 2012; Byrd et al., 2015; Coalson & Byrd, 2017). Further, kinematic differences, when measured, often show robust differences even at lower loads in CWS and AWS (Sasisekaran, 2013; Sasisekaran & Weisberg, 2014; Smith et al., 2010). Differences in kinematics in fluent nonword repetition have been taken as evidence of inefficient, immature, or compromised motor-control processes (Smith et al., 2010) or as compensatory motor control processes in AWS that are related to compromised speech motor skill acquisition (Namasiyavam & van Lieshout, 2008). It is also important to consider that,
because differences in speech kinematics are present during perceptually fluent nonword repetition in AWS, differences coordination,
amplitude, and slower adaptation over trials may also be interpreted as momentary compensation for instabilities in speech motor
control as a function of nonword repetition load.

Despite variability in nonword repetition findings across studies, there are a number of common themes. Interestingly, one
common finding across recent studies is that preschool CWS tend to present with lower accuracy on nonword repetition tasks but do
not present with differences on verbal repetition tasks with semantic content, suggesting a dissociation between verbal working
memory and PWM early in DS (Pelczarski & Yaruss, 2016; Spencer & Weber-Fox, 2014). One possible interpretation is that the
dissociation may be related to compensatory processing in Baddeley’s EB for words with semantic content that is not available for
nonwords, suggesting that LTM for language may play a compensatory role (Pelczarski & Yaruss, 2016). As noted in Spencer and
Weber-Fox (2014), successful nonword repetition requires at least intact auditory-phonological processing, phonological analysis and
storage, speech planning, and ultimately execution. Those authors suggest that the amalgam of those processing demands may index
a fundamental process predictive of recovery and persistence. Finally, given the results of Anderson and Wagovich (2010), it also
seems likely that some aspects of attention also interact with those processes in preschool CWS, suggesting that aspects of the CE may
also be involved in observed differences. Further, differences in school-age CWS and AWS have only been observed for higher
nonword repetition loads, suggesting that aspects of executive function/attention may play an important role in observed differences.
Differences in each process, including the processing of speech sounds (Neef et al., 2012; Saltuklaroglu et al., 2017), phonological
encoding (Kolk & Postma, 1997), executive function/attention (Alm, 2004; Alm & Risberg, 2007; Eggers et al., 2012), and differences
in linguistic planning and execution (Howell & Au-Yeung, 2002; Kleinow & Smith, 2000), have been proposed to play some role in DS
(Bloodstein & Bernstein-Ratner, 2008). As such, while nonword repetition appears to index an important process related to persist-
ence/recovery, the underlying mechanisms relating nonword repetition to stuttering are unclear.

2.2. A review of the evidence: dual-task performance in DS

Given that the processes of attention and executive function likely interact with the function of the PL, consideration of those
capacities is critical when discussing the potential role of PWM in DS. Behavioral evidence for the function of the CE in Baddeley’s
model is often examined through the lens of dual-task experimental paradigms (Baddeley, 2012). In these dual-task experiments,
participants complete mental operations such as counting backwards while their performance is evaluated on some target task.
Baddeley and colleagues suggest that such tasks interrupt the automaticity of the target task, placing load on aspects of cognition
subserved by the CE component (Baddeley, 2012). Following that reasoning, dual-task experiments could potentially provide a
window into the link between PWM and aspects of cognition relegated to the CE in PWS. Dual-task studies in AWS have investigated
aspects of target and secondary task performance as a measure of processing load along with effects on measures of stuttering.
As such, in the sections below we review evidence for differences in behavioral performance in dual-tasks along with evidence sug-
gestng effects on stutter-typical disfluencies.

2.2.1. Dual-task performance in school-age CWS and AWS

While a number of early dual-task studies investigated the performance of distracting tasks on aspects of speech and language
production in school-age CWS and AWS, those studies are difficult to interpret because it is unclear what cognitive processes were
engaged by secondary tasks. Early studies employed secondary tasks such as flickering lights, tracking a moving dot pattern, finger
tapping, gross motor activity, and tracking line patterns while CWS and AWS were engaged in various speech production tasks
(Arends et al., 1988; Brutten & Trotter, 1985; Brutten & Trotter, 1986). Two of the studies were conducted with school-age CWS and
teenage CWS (age range 6–17 years) and reported somewhat lower finger-tapping rates as verbal task demands increased (Brutten
& Trotter, 1985; Brutten & Trotter, 1986). Subsequent studies in AWS have employed secondary tasks that were more readily inter-
pretable. In one early investigation, Caruso and Chodzko-Zajko (1994) reported differences in the accuracy of a speaking (speech
response latency and rate) while AWS and matched TFS performed the Stroop color word task. More recently, Bosshardt and col-
leagues conducted a series of investigations in which AWS and matched TFS performed various speaking tasks along with a number of
secondary tasks including mental calculation, rhyme and semantic category decisions, and memorizing phonologically similar and
dissimilar words (Bosshardt, 2002; De Nil & Bosshardt, 2006). In the speaking tasks, group differences or interactions between task
and group showed differences in measures of inhalation rate or word duration. In sentence level speaking tasks, decreases in the
number of propositions were also found in dual-task conditions relative to single task conditions (Bosshardt, 2002), suggesting that
AWS responded to task demands by reducing speech/linguistic output relative to TFS. In a review of those studies, Bosshardt (2006)
concluded that AWS compensate for the effects of processing load in dual-tasks by reducing the ‘conceptual work’ of speech and
language production in an attempt to decrease stuttering.

More recently, Smits-Bandstra and De Nil (2009) compared a group of AWS and TFS on repeated practice of ten syllable nonsense
sequences while performing a color recognition task. AWS demonstrated slowed speech sequence initiation over practice compared to
the TFS group and decreases in syllable sequence reading accuracy. Jones, Fox, and Jacewicz (2012) used a dual-task in which
participants made rhyme judgements while performing a letter recall task. While rhyme judgements did not show group differences,
reaction time, and letter recall decreases in accuracy for higher loads (i.e., five letters) were observed. The differences in response
time as opposed to accuracy in rhyme judgements suggest that AWS may need more time to process phonological information.
Consistent with nonword repetition findings, apparent behavioral differences were only detected when phonological processing was
heavily loaded. In accord with behavioral findings, a recent electrophysiological study also reported group differences in event-
related potentials (ERPs) in a dual-task wherein participants performed a tone discrimination task while at the same time naming
pictures (Maxfield et al., 2016). Whereas no differences in ERPs to tones were reported in the simple tone judgement task, ERPs were reduced in the AWS group during the dual-tasks (i.e., phonological competition and lexico-semantic conditions). Most recently, Eichorn, Marton, Schwartz, and Melara (2016) reported no group differences in behavioral performance on a speaking task (narrative retell) performed along with a range of tasks designed to load working memory/attention, including a digit span and a spatial working memory task. While no differences in behavioral performance were found, the AWS group did reduce the number of syllables produced in the dual-tasks, suggesting that AWS adjusted speech and language to accommodate processing load. Taken together, the majority of findings support the notion that AWS make subtle adjustments to speech and language production to accommodate processing limitations in dual-task compared to single task conditions.

2.2.2. Dual-task effects on stuttering

While many dual-task studies have reported subtle adjustments in speech or language, a smaller subset have also reported effects on stuttering in dual-tasks relative to speaking tasks alone. Early studies employing a range of secondary tasks largely unrelated to speech production or linguistic processing, found no effect on fluency or more often decreases in stuttering under dual-task conditions (Bajaj, 2007). By contrast, in one early study Caruso and Chodzko-Zajko (1994) reported increases in measures of stuttering in the Stroop color-word task. Bosshardt and colleagues reported a complex relationship between dual-task conditions and measures of stuttered speech in which secondary tasks that have minimal overlap with speech and language (e.g., mental calculation) result in reductions in stuttering (Bosshardt, 1999), while tasks that have ostensibly higher overlap with speech and language (e.g., memorizing phonologically similar or dissimilar words) are associated with increases in stuttering (Bosshardt, 2002). Other conditions in which AWS reduce speech and language output in a sentence generation task along with a secondary task (i.e., rhyme and category decisions) were associated with no significant change in measures of stuttering in dual relative to single task conditions, presumably because language output was reduced (Bosshardt, 2002). In contrast with those findings, recent evidence has also demonstrated decreases in measures of stuttering in dual-task relative to single task conditions. Vasic and Wijnen (2005) reported that stuttered speech reliably decreased in a dual-task in which AWS played video games while speaking. Similarly, Eichorn et al. (2016) also reported decreases in measures of stuttering under dual-task digit-span and spatial-working memory tasks. Taken as a whole, the available evidence suggests that the cognitive processes loaded in dual-task conditions may tap into a cognitive resource that is available to adjust the speech and language output in between dual tasks.

2.2.3. Implications of dual-task findings

Collectively, findings in AWS implicate behavioral adjustments to speech and language output in dual relative to single task conditions. Despite findings suggesting dual-task adjustments to speech there are a number of unresolved issues. First, although the evidence implicates differences in behavioral performance between AWS and controls on a range of dual-tasks, there is much less evidence for differences in CWS (Anderson & Wagovich, 2010; Brutten & Trotter, 1986). Although not a dual-task study, Eichorn, Martin, and Pirutinsky (2017) recently reported that school-age CWS have more difficulty with attentional control in a cognitive task measuring cognitive flexibility, suggesting that aspects of executive function may indeed be ‘weak’ in CWS. Second, due in part to the wide array of secondary tasks employed in dual-task studies, it is difficult to tease apart underlying mechanisms mediating differences in behavioral performance or stuttering. While the mechanisms are unclear, a number of explanations have been offered. Smits-Bandstra and DeNil (2007) suggested that reduced dual-task performance is reflective of low levels of automaticity mediated by basal-ganglia-thalamo-cortical (BGTC) loops. In that case, deficits or differences in the BGTC loops of PWS cause decreases in new skill learning that do not become automatic with practice at the same rate as TFS. Under dual-task conditions, attentional load further degrades the process of automation over a number of trials, suggesting a neurophysiological link between attentional resources, automation, and adaptive sensorimotor learning.

In more recent discussions, Maxfield et al. (2016) further propose that attentional resources allocated to atypical linguistic processing related to covert behavior in AWS (e.g., word avoidance) may have significant effects on word retrieval, speech-motor planning, and subsequent execution. Studies reporting greater fluency also make reference to attentional and working memory mechanisms (Eichorn et al., 2016; Vasic & Wijnen, 2005). More specifically, Eichorn et al. (2016) interpret findings of decreased stuttering in dual-task conditions as consistent with the Matched Filter Hypothesis (Chrysikou et al., 2014). According to that framework, top-down attentional control can be detrimental for tasks that require relatively automatic motor performance (i.e., procedural tasks). According to Eichorn et al. (2016), when working memory/attentional resources are otherwise engaged and top-down control of speech is reduced, more automatic mechanisms (e.g., subcortical mechanisms) facilitate speech production. Thus, the effects of top-down attentional control on speech in dual-tasks is consistent with mechanisms mediating enhanced fluency in some contexts (Eichorn et al., 2016) and deleterious effects in others (Bosshardt, 2006; and see Metten et al., 2011), perhaps depending on the direction of attention or the extent to which dual-tasks share cognitive resources for speech and language.

2.2.4. Implications of nonword repetition and dual-task findings

Collectively, nonword repetition and dual-task findings support the notion that limitations on the cognitive capacities underlying PWM may be related to stuttering from the preschool years into adulthood. First, the majority of studies in preschool CWS suggest a load dependent vulnerability to cognitive factors also known to load PWM, including phonological analysis, temporary storage, and speech motor planning prior to nonword repetition. While differences in behavioral accuracy are likely to be subclinical (Pelczarski &
Yaruss, 2016) and may interact with other cognitive-linguistic capacities (Anderson & Wagovich, 2010; Smith et al., 2012), they are also related to subsequent recovery and persistence (Spencer & Weber-Fox, 2014). Further, load-dependent differences are present even in older school-age CWS and AWS. As such, the behavioral evidence from nonword repetition paradigms suggest subtle, load dependent limitations on what Baddeley described as the PL in both CWS and AWS relative to TFS. We suggest that this difference in PWM capacity may be of importance because the processing demands required in nonword repetition appear to be related to persistence and recovery (Spencer & Weber-Fox, 2014; Mohan & Weber, 2015) and may interact with other psycholinguistic processes (e.g., syntax) over the course of language development to influence the expression of stuttering (Usler & Weber-Fox, 2015).

Second, while it is unclear from nonword repetition studies how PWM may be related to overt stuttered moments, dual-task studies suggest that loading aspects of attention in speech production is associated with effects on stuttered speech, especially under conditions in which PWM is loaded (Bosshardt, 2002). However, it also appears that effects on stuttering sensitively depend on study conditions, with some studies providing evidence of increases in stuttering (Bosshardt, 2006; Caruso & Chodzko-Zaiko, 1994) and others paradoxically reporting decreases in stuttering (Eichorn et al., 2016; Vasic & Winjen, 2005). Stutter-typical disfluencies may vary based on the internal attention to speech and language production or the extent to which secondary tasks load overlapping processes required for speech and language processing (Bosshardt, 2006). Thus, while the underlying mechanisms mediating dual-task findings are unclear, they do implicate aspects of what Baddeley described as the CE and PL in the modulation of language production and stuttering under high cognitive-load conditions. We suggest that those findings are important because it is likely that AWS also experience greater cognitive-linguistic loads in more naturalistic communicative exchanges when compared to TFS.

3. PWM in neurobiological models of language and cognition

It is important to consider that, if CWS and AWS do present with differences in executive control interacting with a phonological loop, then the neurocognitive networks identified in CWS and AWS should be consistent with neurobiological models proposing mechanisms to account for those same cognitive capacities. Toward bridging the gap between cognitive-behavioral findings and neurobiological models, the sections below describe models of language proposing that PWM is an emergent property of sensory and motor regions within the perisylvian network (Buchsbaum & D’Esposito, 2008; Hickok & Poeppel, 2007; Hickok et al., 2011; Majerus, 2013). In addition, because DS is associated with differences in attentional control/executive function in both CWS and AWS, we also describe a model proposing that interactions between the prefrontal cortex (PFC) and basal ganglia (BG) are also critical for working memory under distracting or attentionally demanding conditions (Hazry et al., 2007). Finally, we review brain imaging findings that implicate differences in large-scale networks supporting executive function/attentional control, PWM, and speech processing in the pathophysiology of both CWS and AWS. Taken as a whole, we suggest the networks implicated in the pathophysiology of stuttering are consistent with load dependent differences in behavioral accuracy in nonword repetition and dual-tasks.

3.1. PWM in the dual-stream model of language

While DS is most often associated with functional and structural differences within perisylvian regions (Etchell et al., 2018), it has only been in the last decade that neurobiological frameworks have proposed that perisylvian regions support PWM (Buchsbaum & D’Esposito, 2008; Herman et al., 2013; Jacquemot & Scott, 2006; Majerus, 2013). In an early model, Jacquemot and Scott (2006) proposed that auditory-sensory regions and premotor regions act as cyclic input and output buffers respectively allowing for phonological storage, suggesting a link between sensorimotor mechanisms in natural speech production and those involved in PWM. More recently, Hickok and Poeppel’s (2007) ‘dual-stream’ model of speech and language processing also proposed that interactions between premotor, auditory-sensory, and audiomotor regions account for the emergent properties of PWM. According to their dual-stream model, a left hemisphere dominant dorsal stream, including aspects of the posterior inferior frontal gyrus (IFG), premotor cortex (PMC), and the planum temporale (PT) at the parieto-temporal boundary (Spt), mediate the conversion of input from auditory-phonological regions to articulatory output in the IFG/PMC regions. In the model, the primary auditory cortex along the superior temporal gyrus (STG) is involved in spectrotemporal processing bilaterally and a left dominant region along the superior temporal sulcus (STS) is involved in retrieving phonological codes. A bilateral ventral stream in the anterior and more posterior aspects of the middle temporal gyrus (MTG) and inferior temporal sulcus are more heavily involved in processing receptive language and mediate lexical access (see Hickok & Poeppel, 2007 for visual representation). Finally, although the dorsal and ventral streams are proposed to interact in language formulation and processing, it is as yet unclear how and when the two streams interact to allow for speech and language production in naturalistic contexts.

In the dual-stream model, short-term phonological memory (hereafter referred to as PWM) is proposed to emerge from interactions between auditory-phonological regions in the STS, audiomotor mapping in the PT, and premotor regions of the dorsal stream responsible for motor programming and execution (Buchsbaum & D’Esposito, 2008; Hickok et al., 2003). In particular, a region of the PT at the sylvian parieto-temporal junction (i.e., area Spt) is active during tasks in which covert rehearsal is required for maintenance of phonological information in working memory. The same region is active for auditory encoding, subvocal rehearsal, and overt production in a number of tasks, including multisyllabic nonword repetition (Buchsbaum et al., 2001), ‘jabberwocky’ phrase repetition, and even covert rehearsal of piano melodies (Hickok et al., 2003). Finally, the dorsal stream network also mediates behavioral performance on common clinical nonword auditory processing tasks, even in the absence of a requirement for production, strongly suggesting that the dorsal stream mediates PWM even when execution is not required (Perrachione et al., 2017). Thus, the dual-stream model suggests that the integration of sensory and motor representations are critical both for speech production and the
emergent properties of PWM. From this point of view, damage or disrupted input to the left hemisphere dorsal stream would be expected to cause both effects on speech production and tasks that load PWM.

3.2. Disorders of dorsal stream sensorimotor integration and PWM

Working from a dual-stream model of language and a state feedback motor control (SFC) model, Hickok et al. (2011) proposed that both DS and an acquired language disorder known as conduction aphasia (CA) are consistent with dorsal stream disruptions affecting speech production. From that point of view, individuals with CA have particular difficulty with word repetition related to impaired PWM caused by damage to area Spt (Buchsbaum et al., 2011; Hickok et al., 2011). For individuals with CA, phonological errors in natural speech production occur relatively infrequently but are more likely when producing multisyllabic and low frequency words, suggesting parallels between the loci of phonological errors in CA (Buchsbaum & D'Esposito 2008; Buchsbaum et al., 2011) and similar linguistic factors known to influence the likelihood of stutter-typical disfluencies in DS (Howell & Au-Yeung, 1995, Howell & Au-Yeung, 2002). Thus, in the same way that individuals with CA experience load-dependent phonological errors in speech production and poor word repetition, PWS experience load-dependent disfluencies and load-dependent effects in nonword repetition tasks.

According to an SFC model of speech production, for accurate speech production to occur, an inhibitory model of the intended sensory consequences (i.e., a forward model) is generated in premotor and motor regions (i.e., a motor-phonological system), while an excitatory, corrective model is generated in somatosensory and auditory regions (i.e., a sensory-phonological system) (Hickok et al., 2011; Houde & Nagarajan, 2011). An audiomotor translator in area Spt mediates the conversion of motor and sensory phonological information. When the inhibitory forward model is accurate, no correction is needed and the two models are considered to cancel out. However, in instances wherein an error in the forward model is detected, excitatory corrective models are sent back to the PMC and primary motor areas for correction prior to overt production (Hickok, 2012). In people with CA, damage to Spt is proposed to disrupt forward predictions sent from motor-phonological areas to the auditory-phonological regions, causing errors in the internal monitoring process. However, because the auditory-phonological regions are intact, a person with CA is nonetheless able to detect the error and engages in repeated and often unsuccessful attempts to correct it.

By contrast, in DS the motor-phonological and sensory-phonological systems are intact but the audio-motor translator provides ‘noisy’ or high variance input to the auditory-phonological system (Hickok et al., 2011). Due to this noisy input, forward sensory predictions generate estimates of vocal-tract states and auditory-phonological targets that cause an invalid error signal to be sent back to the motor-phonological system. The behavioral consequence of this noisy translation is a predict-correct loop that, under conditions of high sensorimotor ‘load,’ results in overt disfluencies (e.g., part-word repetitions). Further, Hickok et al. (2011) predicted that errors in sensorimotor mapping should be proportional to the load on the network. From this point of view, higher sensorimotor ‘load’ would be expected on low frequency words, multisyllabic words, or potentially in situations taxing other systems related to speech such as when speaking under time-pressure. Because nonword repetition tasks rely heavily on the same network, one prediction of this approach would be that high variance sensorimotor integration would also result in load dependent errors in repetition as reported in both preschool CWs at lower loads (Pelczarski & Yaruss, 2016; Spencer & Weber-Fox, 2014) and older CWs and AWS at higher loads (Byrd et al., 2012; Byrd et al., 2015; Coalson & Byrd, 2017; Sasinikaran, 2013). As such, the above described mechanisms provide a potential neurophysiological link between internal models of motor control proposed to explain the emergence stutter-typical disfluencies (Brown et al., 2005; Guenther et al., 2006; Hickok et al., 2011; Max et al., 2004; Tian & Poeppel, 2012) and the dorsal sensorimotor system proposed to support PWM for learning new words, new phonological representations, and errors in natural, connected speech production (Buchsbaum & D’Esposito, 2008; Hickok et al., 2011; Jacquemot & Scott, 2006; and see Page et al., 2007).

3.3. PWM in neurobiological models of executive control

While a dorsal stream deficit in DS is plausible, it is unclear from a dual-stream model alone how attention and executive function under high load conditions might interact with the dorsal stream. Whereas Baddeley’s cognitive model proposes interactions between the CE and PL, other models suggest that task relevant information is not maintained in dedicated storage areas but as a subset of processes that are in the focus of attention (Cowan, 1999; Cowan et al., 2005). A number of neurobiological accounts have implicated the dorsolateral prefrontal cortex (BA 46/BA 9) and ventrolateral PFC regions (BA 47/45) in attentional capacities along with the inferior parietal lobule (IPL) and superior parietal lobule (SPL) (e.g., Baddeley, 2003; Buchsbaum & D’Esposito, 2008; Majerus, 2013; Nee et al., 2013). Evidence from neuroimaging studies examining working memory in the contexts of visual, auditory, and affective stimuli also suggest that PFC regions interact with other networks depending on the modality of the stimulus to be held in working memory (e.g., visual vs. auditory). In this way, the same networks involved in sensorimotor representations of the object to be recalled maintain it via interactions with attentional and executive control regions in the PFC (D’Esposito, 2007).

Other neurocomputational approaches emphasize interactions between the PFC and BG in working memory performance under attentionally demanding or distracting conditions (Baier et al., 2010; McNab & Klingberg, 2008; Voytek & Knight, 2010). According to the prefrontal cortex and basal ganglia model of working memory (PBWM), the BG provides a ‘go’ or ‘no-go’ signal to the PFC to modulate information relevance while more primitive BG to premotor connections (e.g. supplemental motor area and anterior cingulate cortex) modulate action selection and execution (Hazy et al., 2007; Mink, 1996). In the PBWM computational model, dynamic gating via midbrain structures is learned via reinforcement mechanisms thought to rely on dopaminergic systems, while PFC representations are updated using Hebbian and error-driven learning mechanisms. Bidirectional excitatory connections from the
Auditory feedback control. The results suggested that when accumulating errors between the feedforward and feedback models were
articulator space (DIVA) computational model to simulate an inefficient feedforward model, forcing the model to rely on time-lagged
the phonological errors observed in CA. There is not strong support for the notion that phonological errors of the kind observed in CA
For example, it is unclear how a difference in dorsal stream processing might lead to stutter-typical disfluencies in PWS as opposed to
3.4. PFC, BG, and dorsal stream sensorimotor timing

While interactions between executive control in the PFC, BG, and dorsal regions involved in phonological ‘storage’ appear to play
a role in DS, one potential problem is that the mechanism underlying ‘noisy’ sensorimotor integration is somewhat underspecified.
For example, it is unclear how a difference in dorsal stream processing might lead to stutter-typical disfluencies in PWS as opposed to
the phonological errors observed in CA. There is not strong support for the notion that phonological errors of the kind observed in CA
covary with stuttering (Nippold, 2002; Nippold, 2012), suggesting that the sensorimotor errors in PWS are likely to involve a different
process. One possibility is that, rather than phonological errors caused by damage to the Spt region as in CA, disrupted sensorimotor
integration is instead a matter of inaccurate coordinative timing mediated by the dorsal sensorimotor stream and connections with
subcortical structures also implicated in speech-motor timing (e.g., the BG and cerebellum). While findings in AWS regarding timing
in non-speech motor control have been variable, the notion that stuttering is primarily a disorder of timing or coordination for speech
production has long been the subject of theoretical proposals (Alm, 2004; Etchell et al., 2014; Kent, 1984; Perkins et al., 1991; Smith
& Weber, 2017; Van Riper, 1982). As such, subtle timing differences in sensorimotor integration could potentially account for the
process disrupting both fluent speech production in PWS and differences in tasks loading PWM. In the sections below, we describe
neurobiological frameworks and findings implicating speech-motor timing networks in sensorimotor integration and internal models
of motor control.

3.4.1. Neurobiological frameworks supporting sensorimotor timing differences in DS

Similar to the dorsal stream SFC model described in Section 3.2, connections between the BG and frontotemporal-parietal net-
works involved in sensorimotor timing for speech have also been used to explain a wide range of phenomena in stuttering (Alm,
2004; Alm & Risberg, 2007; Civier et al., 2013; Etchell et al., 2018; Wu, Maguire, Riley, Lee, Keator, & Tang, 1997). In an influential
proposal, Alm (2004) posited that differences in dopamine receptor concentrations in the BG early in speech motor development
cause sensorimotor instabilities in CWS. According to Alm (2004), unstable timing cues provided by the BG are read out to the
supplemental motor area (SMA) for subsequent movement sequencing, eventually resulting in repetitions or prolongations of
movement to compensate for the deficit. Drawing on Goldberg’s dual-premotor hypothesis, Alm (2004) suggested a medial premotor
system (SMA and BG) is more heavily involved in overlearned, automatic, and internally initiated movements whereas a lateral
premotor system (lateral PMC and cerebellum) is more heavily involved in the control of externally cued, less automatic, and novel
movement patterns. The increased fluency well-known to result from altered auditory feedback, choral speech, and rhythmic speech
(e.g., speaking to a metronome) were proposed to be the result of a shift from a compromised medial premotor system to that of an
uncompromised (or less compromised) lateral premotor system. In addition to fluency enhancing effects, it was also proposed that
attention to compensation for sensorimotor instability may shift internal timing cues from the medial to lateral premotor system
when speech production is less automatic or overlearned (e.g., when using fluency shaping techniques). More recently, Alm (2014)
also proposed that situational variability related to social cognition mediated by the polar medial PFC (BA10), anterior cingulate
cortex (ACC), and SMA could be related to situational variability in stuttering during social interactions. From that point of view, the
medial PFC through connections with the BG and premotor regions may modulate motor control depending on social-cognitive
demands. Thus, connections between the PFC, BG, and frontotemporal-parietal networks are implicated in modulating sensorimotor
control and speech fluency depending upon context.

3.4.2. Neurocomputational modeling evidence supporting a link between BG timing networks and internal models

In addition to PFC and BG timing networks proposed by Alm (2004), elevated dopamine levels are also implicated in interaction
with an overreliance on sensory feedback mechanisms derived from internal models of motor control (Civier et al., 2010; Civier et al.,
2013; Max et al., 2004). In a computational modeling study, Civier et al. (2010) used the directions in auditory space to velocities in
articulator space (DIVA) computational model to simulate an inefficient feedforward model, forcing the model to rely on time-lagged
auditory feedback control. The results suggested that when accumulating errors between the feedforward and feedback models were
large (indexed by formant transitions for the auditory targets), self-repair was initiated in the form of a ‘reset’ or syllable repetition. A second modeling study used the gradient order DIVA model (GODIVA) to simulate two proposed brain differences in AWS, a white-matter anomaly in the ventral motor cortex (vMC) and elevated dopamine levels (Civier et al., 2013). Those authors reported that whereas elevated dopamine levels affecting the BG could account for syllable initial disfluencies, the white-matter anomaly could only account for disfluencies on syllables in the middle or later in an utterance, suggesting that the two potential causes of disfluency likely interact. Computational modeling findings using the GODIVA model have recently been extended to nonword syllable repetition (Markiewicz & Bohland, 2016) with more specific descriptions of the functional role of various input and output buffers in syllable production and their interaction with the BG (see Bohland, Bullock, and Geunther, 2010). Taken together, computational modeling findings do suggest a link between the functional role of the BG in speech production and internal models of motor control.

The potential link between timing cues for speech mediated by the BG and internal model proposals is of interest for explaining the relationship between DS and PWM because timing in internal models of motor-control have also been implicated in PWM (Herman et al., 2013). Similar to neurobiological models of speech motor control, the maintenance of phonological traces between motor-phonological and sensory-phonological regions would be expected to require timely cyclic interactions (Jacquemot & Scott, 2006; Hickok et al., 2011; Herman et al., 2013). That is, cyclic activation and suppression between motor and sensory buffers must be maintained prior to execution both in repetition tasks requiring maintenance of phonological traces and in natural speech production (Herman et al., 2013; Jacquemot & Scott, 2006; and see Page et al., 2007). As such, one likely source of variance in sensorimotor input to the left PT region in PWS is generally slowed or mistimed information for phonological/lexical targets with subsequent effects on execution. Further, it is likely that areas of the PFC modulate sensorimotor timing and dorsal stream coordination depending on context (Alm & Risberg, 2007). To explore that possibility, in sections following we review functional and structural brain imaging evidence implicating connectivity in large-scale neurocognitive networks in CWS and AWS consistent with differences in the connectivity of PFC, BG and dorsal sensorimotor regions implicated both in speech production and PWM.

3.4.3. Functional imaging findings in CWS

One source of evidence for differences in large-scale neurocognitive networks in CWS comes from functional imaging studies. Recent functional magnetic resonance imaging (fMRI) studies implicate differences in activation or connectivity of large-scale networks related to attention/executive function and timing for the sensorimotor control of speech (Etchell et al., 2018). In resting-state recordings of children 3–9 years of age, hypo-connectivity motor (IFG/primary motor cortex/PMC), auditory (MTG/STG), and somatosensory association regions (e.g., the supramarginal gyrus; SMG) have been reported, suggesting that a decreased capacity for timing emerges early in intrinsic BGTC networks. Further, hyperconnectivity in the right superior frontal gyrus and putamen was also found in CWS relative to CWNS, implicating connections between the PFC and BG (Chang & Zhu, 2013). Another recent study of cortical connectivity with the BG (i.e., the putamen) during rhythm perception also found decreased connectivity with the left PMC, left STG, SMA, and cerebellum that was also related to lower behavioral performance in CWS relative to CWNS (Chang, Chow, Weiland, & McAuley, 2016). In a study of resting-state cerebral blood flow of CWS and AWS, the CWS were reported to have lower perfusion in Broca’s area and the superior frontal gyrus relative to TFS that inversely correlated with stuttering severity (Desai et al., 2016).

Most recently, differences in metabolite ratios between school-age and adolescent CWS and TFS (age range 5–17 years) within wide-spread executive function/attentional networks, speech sensorimotor control networks, and networks involved in memory were also reported using proton chemical shift imaging (O’Neill et al., 2017). Finally, a recent longitudinal study in CWS reported anomalous connectivity in the resting-state between what is known as the default mode network (DMN), dorsal attention network (DAN), ventral attention network (VAN), fronto-parietal network (FPN) (including the PT region), and somatomotor network, further implicating broader attentional, audiomotor, and somatomotor regions in the pathophysiology of stuttering early in speech and language development (Chang et al., 2017). More importantly, anomalous connectivity between the DMN, DAN, VAN, and fronto-parietal network also predicted later persistence in CWS (Chang et al., 2017). Thus, taken together, functional neuroimaging findings implicate large-scale neurocognitive networks involved in attention, speech-motor timing, and sensorimotor integration early in the preschool and school-age years. Because the cognitive capacities implicated in those networks are also related to working memory brain activation and behavioral performance (Zhu et al., 2013), those findings are consistent with differences in neurocognitive networks supporting PWM reported in the behavioral literature (Spencer & Weber-Fox, 2014).

3.4.4. Structural imaging findings in CWS

In addition to functional imaging studies, a relatively small number of studies have also investigated structural differences between CWS and CWNS. Those studies also implicate anomalies in regions considered to be important for attention/executive function, sensorimotor integration, and PWM. Studies of gray matter volume (GMV) have reported reduced GMV in the putamen and medial frontal gyrus (MFG), whereas the right STG is reported to have higher GMV in CWS ages 9–12 years (Chang, Erikson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008) and CWS ages 6–12 years (Beal et al., 2015). A comparison of recovered and persistent school-age CWS also showed more GMV the left and right STG in children who persist, suggesting that temporal lobe auditory sensory regions play an important role in persistence and recovery (Chang et al., 2008). Studies of gray matter morphology in CWS have also reported differences in the laterality of prefrontal regions in school-age children (8–13) (Mock et al., 2012) and the caudate nucleus of the BG (Foundas et al., 2013). In addition to gray matter differences, weaker white matter connections (as measured by fractional anisotropy; FA) between the left putamen and cortical networks associated with both cognition and motor control (MFG, SMA, left IFG and insula) have also been observed along with decreased connectivity between the left IFG and pSTG (Chang & Zhu, 2013; Chang et al., 2015). Dorsal stream white matter connections, including the left IFG and pSTG are also correlated with...
disassociations on a range of neuromotor, articulatory, and receptive/expressive language measures, suggesting that structural differences may underlie decreased neural efficiency for speech and language production (Choo et al., 2016). As such, structural findings also implicate networks thought to support cognitive capacities underlying working memory emerging early in the development of stuttering (Zhu et al., 2013)

### 3.4.5. Functional neuroimaging findings in AWS

While more than two decades of neuroimaging studies have shown widespread functional and structural differences in AWS relative to TFS, findings have also been inconsistent depending in part on the tasks used, sample sizes, neuroimaging approach and analysis methodology (see Etchell et al., 2018 for comprehensive review; Ingham et al., 2012; Neef et al., 2015). Those caveats notwithstanding, there are several lines of evidence implicating differences in the connections between the BG and sensorimotor networks involved in speech production. Neuroimaging studies have reported differences in the BG and along with premotor, auditory, and cerebellar regions compared to TFS in the eyes-open resting state (Civier et al., 2013; Ingham et al., 2012; Ingham et al., 2004; Wu et al., 1995) during speech production tasks (Giraud et al., 2008; Ingham et al., 2012; Ingham et al., 2004; Watkins et al., 2008; Wu et al., 1995), and even during non-speech movements (Chang et al., 2009). Activity in the BG is also associated with measures of stuttering severity and is related to the fluency enhancing effects of treatment (Giraud et al., 2008; Toyomura et al., 2015). Taken together, those individual reports implicate BG and sensorimotor networks in a range of conditions and suggest a relationship with treatment effects.

#### 3.4.5.1. Evidence from meta-analyses

In addition to individual reports, meta-analyses of positron emission tomography (PET) and fMRI studies also implicate structures within the BG with sensorimotor regions involved in speech production (Belyk et al., 2015; Brown et al., 2005; Neef et al., 2015). An early activation likelihood estimation meta-analysis across speech production tasks found that AWS consistently present with increased activation in the right IFG/insula and vermal region of the cerebellum, with decreased activation in the bilateral STG during speech production (Brown et al., 2005). Those findings were interpreted as an error in the process of internal model formulation in which suppression of the auditory areas below baseline (i.e., inhibitory forward models) causes an error signal to be sent back to motor areas that, in turn, causes a predict-correct loop (e.g., part-word repetitions) (Brown et al., 2005). Two more recent functional imaging meta-analyses also reported patterns of activity related to trait (i.e., TFS compared to AWS) and state (i.e., disfluent speech relative to fluent speech) aspects of stuttering (Belyk et al., 2015; Budde et al., 2014). In trait analyses, right-hemisphere over activations include the PMC, primary motor-cortex, rolndic operculum, insula, posterior IFG, orbital IFG, pre-SMA, MFG, IPL, and SPL. Left-hemisphere under activations include laryngeal motor cortex, MTG, STS, vermal region of the cerebellum, and red nucleus of the midbrain, suggesting widespread neurocognitive networks are involved in stuttering by adulthood. By contrast, state analyses comparing disfluent and fluent speech have found areas related to disfluency are primarily in the left-hemisphere (e.g, IFG, SMA, somatosensory cortex, and BG), while areas related to fluency are found in right-hemisphere regions (primary auditory cortex (A1), PT, pSTG, SMG, IPL, & IFG). In those meta analyses, the only fluency related region in the left was the SMG, suggesting that left somatosensory association areas may also play a role in fluency or compensation for sensorimotor instabilities in the audio-motor network (Belyk et al., 2015; Budde et al., 2014; Neef et al., 2015). Taken together, state analyses suggest that regions consistent with typically left-hemisphere dominant frontotemporal-parietal network are under-active in AWS during stuttered speech. On the other hand, fluency related areas in the right hemisphere are hyperactive during fluent speech, suggesting that right-hemisphere structures may aid in compensation for sensorimotor instabilities on the left. Those findings are consistent with differences in dorsal stream regions proposed to mediate the emergent properties of PWM and sensorimotor integration in speech production.

#### 3.4.5.2. Evidence from resting-state networks

Another line of evidence implicating broader neurocognitive networks in DS is recent evidence from the resting state. One distinct weakness of functional neuroimaging studies employing various speaking and reading tasks is that it is difficult to compare activations across studies due to task differences (Chang et al., 2017; Etchell et al., 2018). While early findings suggested a lack of consistent differences in the resting-state, more recent studies using larger sample sizes and more recent analytic approaches have noted differences (Etchell et al., 2018). Xuan et al. (2012) reported that networks associated with cognitive control including increased connectivity in the prefrontal cortex and DMN and increased connectivity in motor and auditory regions. Lu et al. (2012) reported decreases in functional connectivity during resting state in the left IFG and dorsolateral prefrontal cortex compared to TFS. In that study, AWS who received treatment showed reduction in the left IFG and increases in the SMA and cerebellum, suggesting that treatment effects may have been related to compensatory speech timing networks. Yang, Fanlu, Siok, and Tan (2018) also reported reduced resting-state connectivity between the right SMA and BG, along with decreased connectivity between the STG and BG. More recent studies have also found differences in resting state connectivity between the left IFG and primary auditory cortex/Heschl’s gyrus that is correlated with performance on a speech perception task, further suggesting a relationship between intrinsic differences in sensorimotor networks and challenging speech processing tasks (Halag-Milo et al., 2016; Lu et al., 2016). Thus, resting state findings imply differences in large-scale neurocognitive networks in AWS, including aspects of the PFC, BG, and regions involved in dorsal stream sensorimotor integration in speech production and challenging speech perception tasks.

#### 3.4.5.3. Evidence for the PFC in compensation and unassisted recovery

While resting state findings implicate attention and executive function in AWS, they do not explicitly link it to moment-to-moment variability in stutter-typical disfluencies during speech production. However, recent functional neuroimaging findings in AWS do implicate the left and right orbitofrontal cortex (OFC) in compensation for stuttering (Sītek et al., 2016) and in unassisted recovery from stuttering in adulthood (Kell et al., 2017). Using resting state and diffusion MRI, Sītek et al. (2016) reported that AWS with the least severe symptoms as determined by the Stuttering Severity Instrument – Fourth Edition (SSI-4; Riley, 2009) presented with increased functional and structural connectivity between the OFC and cerebellum compared to AWS with higher SSI-4 severity ratings, suggesting that the OFC/cerebellum could be involved in compensatory mechanisms in the more fluent AWS.
Further, long-lasting recovery without treatment (i.e., recovered stuttering) in AWS has been associated with differences in the left ventrolateral prefrontal cortex (orbital IFG; BA47/12) (Kell et al., 2009). More recent functional connectivity analyses have reported ‘hyperconnectivity’ between BA47, BA44, the SMA, and SMG prior to fluency shaping therapy. Following therapy, reductions in connectivity were observed between BA44 (i.e., opercular IFG) and the SMG along with increased coupling between the anterior STG and ventral motor cortex. By contrast, in recovered stuttering, there was decreased connectivity from BA47 to the SMA, SMG, and superior cerebellum, suggesting that the OFC modulates somatomotor connectivity in recovered stuttering differently than in treated AWS. Thus, recent evidence implicates the PFC in both enhancing fluency and potentially in increasing stutter-typical disfluencies, similar to dual-task findings that load aspects of attention and executive function.

3.4.6. Structural findings in AWS

Evidence from recent structural neuroimaging studies comparing AWS and TFS implicates connectivity between PFC, BG, and dorsal stream regions. One line of evidence in particular implicates structural differences in the PT with at least some AWS demonstrating greater overall area and a more symmetric pattern in the PT (Foundas et al., 2001). Enhanced fluency induced by delayed auditory feedback in a subgroup of AWS is also associated with rightward PT asymmetry (Foundas et al., 2004). Other studies have reported increased white matter underlying the right PT, further suggesting that compensatory processing in the right PT may be related to white matter plasticity and compensation for anomalies in the left PT (Jäncke et al., 2004). Despite these early findings, the role of PT asymmetry in DS is uncertain, as later studies with larger sample sizes and more stringent controls (e.g., inclusion of only right-handed adults) have not replicated gray matter PT asymmetry in AWS relative to TFS (Chang et al., 2008; Cykowski et al., 2008; Gough et al., 2018). One possible mediating factor is variability in the age of study participants. For example, Gough et al. (2018) reported a positive correlation between PT asymmetry and age in TFS while no such relationship was observed for AWS. Another possibility is that there are subtypes of AWS in whom PT asymmetry interacts with structural or functional anomalies in other cortical or subcortical structures (Foundas et al., 2013; Gough et al., 2018). Those uncertainties notwithstanding, PT asymmetry or PT connectivity with other dorsal stream structures implicates structural differences in region thought to be critical for audiomotor integration and PWM in the dual-stream model.

A second line of evidence supporting dorsal stream pathophysiology comes from structural differences in regions thought to be connected to the PT. Decreases in gray matter volume of AWS have also been reported in other structures consistent with PWM buffers thought to be important for articulatory rehearsal. These areas include the IFG and aspects of the more anterior STG (Beal et al., 2007; Song et al., 2007). In particular, decreased gray matter volume in the left IFG has been found to negatively correlate with stuttering severity (Kell et al., 2009). Recent findings also suggest that the right BG and bilateral STG functionally and structurally are more strongly connected than in TFS (Sitek et al., 2016). Sowman et al. (2017) have also reported differences in gray matter volume of the caudate nucleus in AWS, while a number earlier studies reported differences in the right precentral gyrus and STG (Beal et al., 2007; Kikuchi et al., 2011; Song et al., 2007). In addition to gray matter anomalies, studies have also reported white matter differences in areas important for connecting premotor structures with somatosensory and audio-motor regions (Cai et al., 2014). Further, anomalies along white matter tracts connecting dorsal stream regions may also be present in the peripheral nervous system along the corticobulbar and corticospinal tracts, suggesting that dorsal stream anomalies may be related to decreased efficiency of peripheral execution (Connally, Ward, Howell, & Watkins, 2014). A recent ALE meta-analysis in AWS also reported decreased FA in white matter tracts connecting left hemisphere frontal motor structures (IFG, vPMC, primary motor cortex) with temporal structures (pSTG, middle temporal gyrus; MTG) and parietal structures (SMG, primary somatosensory cortex) (Neef et al., 2015).

While methodological limitations make attributing white matter anomalies to the dorsal stream as opposed to the ventral stream uncertain, recent studies have reported a lack of white matter differences in the ventral stream with robust differences in the dorsal stream (see Kronfeld-Duenias et al., 2017 for debate). Finally, in the most recent study implicating the PFC in stuttering, Neef et al. (2017) reported that increased connection strength between the right frontal pole (BA10) and subthalamic nuclei is positively correlated with stuttering severity, while the uncinate fasciculus connecting the right frontal pole to auditory and multisensory regions was negatively correlated with stuttering severity, suggesting that aspects of the frontal pole may be related both to compensatory and inhibitory processes affecting fluent speech production. Taken as a whole, the structural imaging literature suggests anomalous connectivity in large-scale neurocognitive networks including the PFC, BG, and dorsal sensorimotor stream implicated in PWM, while differences in the ventral stream important for lexical access remain uncertain.

3.4.7. Implications of neuroimaging findings in CWS and AWS

Although a number of interpretations have been proposed to account for neuroimaging findings in CWS and AWS, a common theme is that both CWS and AWS present with differences in networks thought to mediate timing in speech-motor control and sensorimotor integration. Despite considerable spatial variability across studies, in a systematic review of the neuroimaging literature, Etchell et al. (2018) concluded that one common finding in both CWS and AWS is decreased activation and reduced gray matter volume in the left IFG and connections with the SMA, temporal lobe regions, and right IFG. Additionally, in the only study to examine anomalous structure and function in CWS and AWS, the IFG (pars opercularis) was also reported to be the only region to show an abnormal developmental trajectory (reduced gray matter thinning), suggesting that gray matter anomalies or underlying white matter anomalies are a strong candidate for the cause of anomalous input to the sensorimotor network in the preschool years (Beal et al., 2015). Interestingly, the triangular and opercular portions of the IFG are also causally related to the transmission of information for timing and sequencing in speech production as opposed to articulatory quality located in more dorsal premotor regions (Long et al., 2016). The commonly reported hyperactivity in the right IFG (Belyk et al., 2015; Brown et al., 2005; Budde et al., 2014), thought to compensate for deficits in the left IFG, is also considered a core structure active across a range of conditions requiring the processing of temporal information in perception and motor control (Weiner and Turkeltaub, 2010). As such, neuroimaging findings suggest that differences in the efficiency of timing networks related to both dorsal stream sensorimotor...
integration and the emergent properties of PWM are different in both CWS and AWS. In this way, the available neuroimaging literature suggests that behavioral differences in accuracy on tasks heavily recruiting those same networks, such as nonword repetition or dual-tasks, are likely in both CWS and AWS.

Finally, in CWS regions within the PFC appear to be related to persistence and recovery in connection with sensorimotor networks (Chang et al., 2017). In AWS, subregions of the PFC in connection with sensorimotor networks also appear to be related to compensation for stuttering (Sitek et al., 2016), treatment effects (Kell et al., 2009), and unassisted recovery (Kell et al., 2017). The most recent study to date to interpret functions of the PFC also suggested that different right hemisphere connections may subserve both fluency enhancing compensatory and detrimental inhibitory processes, suggesting that the right PFC via different pathways may be involved in processes that are both deleterious to and facilitate fluency in AWS (Neef et al., 2017). Those findings are broadly consistent with discrepant dual-task effects on stuttering depending on the experimental conditions (Bosshardt, 2006). Taken together, those findings are consistent with a modulatory role for the PFC that may be related to both compensation for sensorimotor instabilities and conditions that increase the likelihood of stuttering.

4. Linking brain and behavior

Differences in connectivity of the PFC, BG, and dorsal stream sensorimotor regions may provide explanations for common behavioral findings in nonword repetition, auditory-perceptual, and dual-tasks in DS. First, disrupted timing cues from the forward model to the inverse model affecting speech production could explain subtle behavioral differences commonly observed on nonword repetition tasks in preschool CWS (Spencer & Weber-Fox, 2014). If differences in sensorimotor timing or ‘high variance’ input to internal models of motor control account for incorrect nonword repetition trials, the process would not always be expected to result in an incorrect response. In other words, on some trials CWS would be expected to compensate for increased variance or interbuffer timing differences much like they would in natural speech production (i.e., perceptually fluent speech) (c.f. Namasiyavam & van Lieshout, 2011). Such a mechanism might account for statistically significant differences in nonword repetition that are nonetheless within normal limits in preschool CWS (Pelczarski & Yaruss, 2016). In AWS, the proposal is consistent with differences in fine motor control during nonword repetition tasks at lower loads (e.g., 4 syllables), suggesting compensation for sensorimotor instabilities (Smith et al., 2010; Smith et al., 2012), while behavioral differences emerge only when the PWM system is heavily loaded (Byrd et al., 2012; Byrd et al., 2015; Coalson & Byrd, 2017). Finally, while lower performance on nonword repetition tasks is commonly observed in CWS, semantic repetition tasks typically do not result in significantly lower accuracy (Spencer & Weber-Fox, 2014; Anderson & Wigovich, 2010; Pelczarski & Yaruss, 2016) and do not appear to be related to persistence and recovery (Spencer & Weber-Fox, 2014). Interestingly, syntactic phrase structure violations embedded in sentences with semantic content are not associated with processing differences between CWS who recover and those who persist, but similar violations are associated with processing differences (i.e., N400-like response) between the two groups in nonsense ‘jabberwocky’ phrase repetition tasks. Those findings suggest limitations on PWM interacting with syntax may also be related to persistence/recovery while semantic processes alone are not (Usler & Weber-Fox, 2015). Taken together, findings are more consistent with dorsal stream pathophysiology than ventral stream pathophysiology proposed to mediate LTM for semantic content or lexical access (Hickok et al., 2011; and see Majerus, 2013) and may explain why single word, verbal repetition tasks typically do not result in significantly lower behavioral performance or predict persistence/recovery.

Second, a deficit in dorsal stream sensorimotor integration is consistent with lines of evidence suggesting neurobiological and subtle behavioral differences between PWS and controls in challenging auditory perception tasks in the absence of a functional deficit in language comprehension (Chang et al., 2009; Halag-Milo et al., 2016; Neef et al., 2012; Saltuklaroglu et al., 2017; Sato et al., 2011). According to the dual-stream model, the dorsal stream may aid perceptual tasks in which working memory and attention are focused on speech-sound analysis, whereas the ventral stream mediates receptive processing wherein the goal is linguistic comprehension (i.e., a double dissociation) (Hickok & Poeppell, 2007). Forward models of speech production may be adapted in some instances to difficult perceptual tasks (e.g., discrimination in noise), allowing for constraints on expected upcoming speech percepts that may later aid working memory for the subsequent response (e.g., via covert rehearsal) (Hickok et al., 2011). Thus, mistimed or ‘noisy’ sensorimotor integration would be expected to cause subtle differences in auditory-perceptual tasks without functional language comprehension deficits. Given the link between sensorimotor integration in speech perception and production, PWM may also be related to a higher risk of phonological disorders in children who stutter (Ambrose et al., 2015). As such, the evidence implicates differences in the dorsal sensorimotor network that are related to auditory speech processing, phonological storage, and the sensorimotor control of speech production in the absence of functional language comprehension deficits thought to be mediated by the ventral stream.

Third, there is some evidence from preschool-age CWS that aspects of attention interact with the process of phonological storage for nonword repetition (Anderson & Wagovich, 2010) and the neurocognitive networks supporting attention and executive function are implicated in persistence and recovery (Chang et al., 2017). A potential role for the PFC and BG in executive control of both working memory under distracting conditions and speech production under attentionally demanding conditions (e.g., social situations or under time-pressure) is consistent with those findings and recent brain imaging evidence implicating both intrinsic and disflueny related neurocognitive networks in AWS and CWS (Chang et al., 2017; Kell et al., 2017; Sitek et al., 2016). In such case, tasks that ensure attention is directed away from speech production, decoupling of the PFC-BG and dorsal sensorimotor/cerebellar network would be expected along with fluency enhancement (e.g., Eichorn et al., 2017; Vasic & Winjnen, 2005). However, in other tasks in which attention is directed internally to speech production (e.g., Bosshardt, 2002; Caruso & Chodzko-Zajko, 1994) or that load highly overlapping cognitive-linguistic processes (Bosshardt, 2002; Metten et al., 2011), greater PFC-BG coupling with the dorsal stream would be expected along with instabilities in dorsal stream sensorimotor timing. That potential function of PFC-BG modulation of the dorsal stream during internal model formulation is broadly consistent with the OFCs role in compensation for sensorimotor instabilities and disconnection of the OFC from the audiomotor and somatosensory regions in unassisted recovery (Kell et al., 2017). Thus, the PFC may be important in compensating for internal sensorimotor
errors in addition to increasing the probability of stuttering depending upon the circumstances. To investigate those possibilities, future work would need to carefully manipulate the direction of attention in dual-tasks and the extent to which cognitive resources are shared between tasks.

5. Future directions: neural oscillations in PWM

While there appear to be links between the neurobiology of stuttering and PWM, it is still unclear what precise mechanisms mediate that relationship. We suggest that mechanisms underlying differences in sensorimotor control and PWM in CWS and AWS could be further explored by investigating cyclic dorsal stream interbuffer timing differences in nonword repetition and dual-tasks guided by the neurobiological networks reviewed. Fig. 1 depicts cortical networks likely to be related to PWM according to the dual-stream model proposed by Hickok and Poeppel (2007) (blue, green, yellow) and more recent models and data proposing an executive network involved in attentional control (red) (D’Esposito, 2007; Majerus, 2013; Nee et al., 2013). While it is difficult to examine cyclic interbuffer timing in the encoding, maintenance, and production phases of a nonword repetition task when using relatively low temporal resolution imaging approaches (e.g., fMRI), one way to investigate ongoing sensorimotor timing in each phase of the task is via event-related changes in the ongoing time-frequency of the electroencephalography (EEG) and magnetoencephalography (MEG) known as ‘neural oscillations.’ According to a ‘dynamic’ model of oscillatory function, neural oscillations are critical for long-distance interregional communication and timely correlations between the phase of oscillations in prefrontal regions and sensorimotor regions may reflect top-down control of interregional communication in both perception and action-control (Fries, 2005; Siegel et al., 2012). As such, the power and phase of neuronal oscillations has the potential to allow for the study of large scale neuronal dynamics in the various stages of nonword repetition related to stimulus encoding, trace maintenance, and repetition (i.e., execution).

Multifactorial models of stuttering have long predicted that even subtle differences in cognitive capacities affect motor control for speech in DS (Smith, 1999; Starkweather, 2002). In particular, the speech motor skill (SMS) framework predicts that cognitive-sensorimotor loads may destabilize central neural oscillators, resulting in decreased oscillatory entrainment with speech-motor effectors critical for feedback control (Namasivayam & van Lieshout, 2011; van Lieshout et al., 2014; van Lieshout et al., 2004). From that point of view, the repetition of novel phonological sequences and dual-tasks loading cognitive-sensorimotor oscillators may cause load-dependent breakdowns in entrained sensorimotor-effectors oscillations. One likely difference related to sensorimotor timing in AWS and CWS is neural oscillations in the beta band (15–25 Hz) (Etchell et al., 2014). The beta band is a known correlate of motor activation and more recently has been proposed to mediate coordinative timing in audio-motor integration for speech along with lower frequency bands (Morillon & Schroeder, 2015). In CWS differences in the beta band have been found at rest (Özge, Torosm & Çomelekoğlu, 2004) and reversals in the timing of peak beta band power reduction (i.e., suppression) and enhancement in a rhythmic tone perception task has been observed in CWS relative to CWNS (Etchell et al., 2016). In AWS, differences have been found beta band interhemispheric connectivity at rest (Joos et al., 2014), in speech production tasks (Mersov et al., 2017), and in speech perception tasks (Saltuklaroglu et al., 2017). As such, we suggest that the study of neural oscillations in nonword repetition and dual-tasks may provide a means of investigating a much needed link between the sensorimotor mechanisms proposed to mediate stutter-typical disfluencies (Hickok et al., 2011; Tian & Poeppel, 2012) and sensorimotor mechanisms mediating PWM.

Fig. 1. A depiction of networks likely to interact in nonword repetition and dual-tasks based on Hickok and Poeppel’s dual-stream model (Hickok and Poeppel, 2007) (blue, green, yellow) and an attentional/executive network in working memory (see Majerus, 2013 and Nee et al., 2013) (red). Regions in red include the dorsolateral prefrontal cortex (DLPFC), the ventrolateral prefrontal cortex (VLPFC), inferior parietal lobule (IPL) and superior parietal lobule (SPL). The dorsal stream includes the IFG, PMC, and PT/Spt regions (blue) along with the primary auditory cortex (green) and superior temporal sulcus (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
behavioral or mindfulness training approaches focusing on cognitive-linguistic and cognitive-emotional processes (Boyle, 2011; McNab & Klingberg, 2008). As such, a better understanding of the mechanisms underlying PWM has the potential to explain how sensorimotor loads or capacities are related to individual differences in brain networks. In the future, investigations across studies (Etchell et al., 2018; Ingham et al., 2012; Neef et al., 2015) and structural imaging methodology is currently limited. For those reasons, differences in particular regions or tracts should be interpreted cautiously (Etchell et al., 2018). Finally, neural networks that are stable within an individual AWS nonetheless differ significantly between AWS, suggesting significant inter-individual differences in neural networks depending on previous experiences with stuttering (Wymbus et al., 2013). Such a finding is not surprising considering that individual PWS have variable phenomenological experiences (Plexico et al., 2005) and to some extent functional reorganization would be expected to reflect those experiences (Ingham et al., 2012; Wymbus et al., 2013).

Despite the above cautions, further investigation of the link between the neurobiological underpinnings of PWM and stuttering has a number of advantages. First, the study of PWM has the advantage of a rich theoretical history, stable and replicable empirical paradigms, and supporting neurobiological networks that have also been implicated in both CWS and AWS. Second, because much of the neuroimaging literature in stuttering has not been driven by current neurobiological models of language and cognition, the application of models outside the field could potentially facilitate collaboration across laboratories. Finally, if PWM is related to variability in the expression of stuttering, it may aid in explaining interindividual differences in brain networks. The networks implicated here predict the extent to which relevant information is stored and are associated with individual differences in working memory capacity (McNab & Klingberg, 2008). As such, a better understanding of the mechanisms underlying PWM has the potential to explain how sensorimotor loads or capacities are related to individual differences in brain networks. In the future, investigations may continue to aid in identifying CWS who are more likely to persist or recover and AWS that would benefit from cognitive-behavioral or mindfulness training approaches focusing on cognitive-linguistic and cognitive-emotional processes (Boyle, 2011; Menzies et al., 2009).
### Appendix A

#### Table A1
Summary of nonword repetition studies in preschool-age children who stutter.

<table>
<thead>
<tr>
<th>Study</th>
<th>Age (range)</th>
<th>N</th>
<th>Matching</th>
<th>Difference in nonword repetition accuracy</th>
<th>Maximum syllable length</th>
<th>Difference in kinematic measures</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hakim &amp; Ratner (2004)</td>
<td>4;1-8;4</td>
<td>8 CWS; 8 CWNS</td>
<td>age, sex, maternal education</td>
<td>yes</td>
<td>5</td>
<td>NM</td>
<td>d = 1.41</td>
</tr>
<tr>
<td>Anderson et al. (2006)</td>
<td>3;0-5;2</td>
<td>12 CWS; 12 CWNS</td>
<td>age, sex, SES</td>
<td>yes</td>
<td>5</td>
<td>NM</td>
<td>η² = .29</td>
</tr>
<tr>
<td>Anderson &amp; Wagavich (2010)</td>
<td>3;6-5;2</td>
<td>9 CWS; 14 CWNS</td>
<td>none</td>
<td>yes</td>
<td>5</td>
<td>NM</td>
<td>η² = .42</td>
</tr>
<tr>
<td>Smith et al. (2012)</td>
<td>4-6</td>
<td>31 CWS; 22 CWNS</td>
<td>SES</td>
<td>no</td>
<td>4</td>
<td>yes</td>
<td>NR</td>
</tr>
<tr>
<td>Spencer &amp; Weber-Fox (2014)</td>
<td>3;9-5;8</td>
<td>21 CWS-Rec; 9 CWS-Per; 25 CWNS</td>
<td>age, SES, nonverbal reasoning</td>
<td>yes</td>
<td>4</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Pelczarski &amp; Yaruss (2016)</td>
<td>5;5-5;8</td>
<td>11 CWS; 11 CWNS</td>
<td>age, language ability, sex, maternal education</td>
<td>yes</td>
<td>7</td>
<td>NM</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Note. CWS = children who stutter; CWNS = children who do not stutter; CWS-Rec = children who stutter, recovered; CWS = Per = children who stutter who persist; SES = Socioeconomic Status; NM = not measured; NR = not reported.

* Smith et al., (2012) did not report significant differences in overall accuracy of nonword repetition between CWS and CWNS groups, but did report a significant difference for CWS with articulation/language disorders.
# Table A2

Summary of nonword repetition studies in school-age children who stutter.

<table>
<thead>
<tr>
<th>Study</th>
<th>Age (range)</th>
<th>N</th>
<th>Matching</th>
<th>Difference in nonword repetition accuracy</th>
<th>Maximum syllable length</th>
<th>Difference in kinematic measures</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seery et al. (2006)</td>
<td>8;6-12;6</td>
<td>14 CWS; 11 CWNS</td>
<td>None</td>
<td>yes*</td>
<td>5</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Bakhtiar et al. (2007)</td>
<td>5;1-7;10</td>
<td>12 CWS; 12 CWNS</td>
<td>None</td>
<td>no</td>
<td>3</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Chon &amp; Ambrose (2007)</td>
<td>10-13</td>
<td>14 CWS-Rec; 5 CWS-Per</td>
<td>None</td>
<td>no</td>
<td>4</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Weber-Fox et al. (2008)</td>
<td>9;4-13;9</td>
<td>10 CWS; 10 CWNS</td>
<td>age, sex</td>
<td>no**</td>
<td>3</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Oyoun et al. (2010)</td>
<td>5-13</td>
<td>30 CWS; 30 CWS</td>
<td>age, sex</td>
<td>yes***</td>
<td>3</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Sasisekaran &amp; Byrd (2013)</td>
<td>8-15</td>
<td>14 CWS; 14 CWNS</td>
<td>age, sex</td>
<td>no**</td>
<td>4</td>
<td>NM</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Note.* CWS = children who stutter; CWNS = children who do not stutter; CWS-Rec = children who stutter, recovered; CWS = Per = children who stutter who persist; NM = not measured; NR = not reported.

* Seery et al., (2006) reported a significant interaction between group and syllable length and reported significant differences in post-hoc comparisons between CWS and CWNS at 5-syllable length.

** Weber-Fox et al. (2008) reported similar nonword repetition accuracy in both groups that was comparable to 30 CWNS in the same age range in a previous study.

*** Oyoun et al. (2010) also reported a significant correlation between subject age and severity with nonword repetition performance.

**** Sasisekaran & Byrd (2013) further subdivided CWS into younger (7–11) and older groups (11–15) and reported significant differences in a phoneme elision task between younger and older groups and no differences between age matched CWNS groups.
<table>
<thead>
<tr>
<th>Study</th>
<th>$N$</th>
<th>Matching</th>
<th>Difference in nonword repetition accuracy</th>
<th>Maximum syllable length</th>
<th>Difference in kinematic measures</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ludlow et al. (1997)</td>
<td>5 AWS; 7 TFS</td>
<td>none</td>
<td>yes</td>
<td>4</td>
<td>NM</td>
<td>NR</td>
</tr>
<tr>
<td>Namasivayam &amp; van Lieshout (2008)</td>
<td>5 AWS; 5 TFS</td>
<td>age, sex, education level</td>
<td>NM</td>
<td>4’</td>
<td>yes</td>
<td>$d = 1.64$</td>
</tr>
<tr>
<td>Smith et al. (2010)</td>
<td>17 AWS; 17 TFS</td>
<td>age, sex, education level</td>
<td>no</td>
<td>4</td>
<td>yes</td>
<td>NR</td>
</tr>
<tr>
<td>Byrd et al. (2012)</td>
<td>14 AWS; 14 TFS</td>
<td>age, sex</td>
<td>yes</td>
<td>7''</td>
<td>NM</td>
<td>$\eta^2 = .15$</td>
</tr>
<tr>
<td>Sasisekaran (2013)</td>
<td>9 AWS; 9 TFS</td>
<td>age, sex</td>
<td>no</td>
<td>4***</td>
<td>yes</td>
<td>$\eta^2 = .32$</td>
</tr>
<tr>
<td>Sasisekaran &amp; Weisberg (2014)</td>
<td>10 AWS; 10 TFS</td>
<td>age, sex</td>
<td>yes</td>
<td>6</td>
<td>yes</td>
<td>NR</td>
</tr>
<tr>
<td>Byrd et al. (2015)</td>
<td>10 AWS; 10 TFS</td>
<td>age, sex, education level</td>
<td>yes</td>
<td>7</td>
<td>NM</td>
<td>$\eta^2 = .38$</td>
</tr>
<tr>
<td>Coalson &amp; Byrd (2017)</td>
<td>26 AWS; 26 TFS</td>
<td>none</td>
<td>yes</td>
<td>2***</td>
<td>NM</td>
<td>$d = .34$</td>
</tr>
</tbody>
</table>

Note. AWS = adults who stutter; TFS = typical fluent speakers; NM = not measured; NR = not reported.

* Namasivayam & van Lieshout (2008) reported practice effects on two bisyllable, bilabial nonword productions. The effect size reported is for a main effect of movement amplitude.

** Byrd et al., (2012) reported a significant interaction between group and syllable length in a nonword condition with 7 syllables.

*** Sasisekaran (2013) reported no differences on a nonword repetition task but did reported differences on kinematic measures during a nonword reading task. The effect size reported is for lip aperture variability.

**** Coalson & Byrd (2017) used only two syllable lengths but manipulated stress patterns and reported a significant interaction of talker group, stress, and task indicating that group differences were driven by only the most demanding conditions (e.g., low frequency stress patterns).
Table A4
Summary of dual-task studies in adults who stutter.

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Matching Differences in primary task/abilities</th>
<th>Differences in secondary task/abilities</th>
<th>Effects on stuttering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caruso et al. (1994)</td>
<td>9 AWS; 9 TFS</td>
<td>none</td>
<td>yes (latency, speech rate)</td>
<td>no (Stroop color word) increases</td>
</tr>
<tr>
<td>Bosshardt et al. (1999)</td>
<td>9 AWS; 10 TFS</td>
<td>mental calculation performance</td>
<td>yes (inhalation rate, length of breath group, word duration)</td>
<td>no (mental calculation) increases primarily in a subgroup of AWS</td>
</tr>
<tr>
<td>De Nil &amp; Bosshardt (2000)</td>
<td>12 AWS; 12 TFS</td>
<td>none</td>
<td>(sentence planning)</td>
<td>‘NR (rhyme and category decisions)’ ‘NM</td>
</tr>
<tr>
<td>Bosshardt et al. (2002)</td>
<td>14 AWS; 16 TFS</td>
<td>age, education, sex, memory span, vocabulary</td>
<td>no (pause rate, inhalation rate)</td>
<td>‘yes (reading and memorizing phonologically similar and dissimilar words)’ increases</td>
</tr>
<tr>
<td>Bosshardt et al. (2002)</td>
<td>14 AWS; 16 TFS</td>
<td>age, education, sex, digit span, vocabulary</td>
<td>*** yes (# of propositions in sentence production)</td>
<td>no (rhyme and category decisions) no change</td>
</tr>
<tr>
<td>Vasic &amp; Wijnen (2005)</td>
<td>22 AWS; 10 TFS</td>
<td>none</td>
<td>no (narrative retell)</td>
<td>‘***NM decreases’</td>
</tr>
<tr>
<td>Smits-Bandstra &amp; De Nil (2009)</td>
<td>9 AWS; 9 TFS</td>
<td>sex, educational background</td>
<td>no (rhyme judgement)</td>
<td>yes (letter recall) NM</td>
</tr>
<tr>
<td>Jones et al. (2012)</td>
<td>15 AWS; 15 TFS</td>
<td>none</td>
<td>yes (naming accuracy)</td>
<td>no (tone judgement accuracy) NM</td>
</tr>
<tr>
<td>Maxfield (2016)</td>
<td>20 AWS; 20 TFS</td>
<td>none</td>
<td>*** no (speech errors)</td>
<td>no (spatial working memory and digit span tasks) decreases</td>
</tr>
</tbody>
</table>

Note. AWS = adults who stutter; TFS = typical fluent speakers; NM = not measured; NR = not reported.

* De Nil & Bosshardt (2000) used a sentence planning task wherein no behavioral measures were available. No measures of language production, stuttering, or statistical tests of rhyming and category decisions were reported.

** Bosshardt et al. (2002) reported significantly lower word recall accuracy in the TFS group compared to the AWS group.

*** Bosshardt et al. (2002) reported significantly a higher percentage of stuttered syllables in the AWS group relative to controls but dual-tasks were not associated with more stuttering than sentence production.

**** Vasic & Wijnen (2005) used a distracting video game in simple and more challenging conditions but did not report behavioral performance on the task.

***** Eichorn et al. (2016) did not report more speech errors in dual-tasks conditions for AWS relative to controls but did report lower speech rate in the AWS group compared to controls.

Andrew Bowers, PhD CCC-SLP is an Assistant Professor at the University of Arkansas where he directs the EEG Neurofeedback Laboratory. He conducts research focusing on sensorimotor integration in speech perception and production using high density electroencephalography (EEG), autonomic measures, and surface electromyography (EMG). His research examines relationships between sensorimotor integration, cognition, and measures of task-performance toward non-invasive brain-computer interface methods designed to augment treatment approaches in fluency disorders.

Lisa Bowers, PhD CCC-SLP is an Assistant Professor at the University of Arkansas. Her research focuses on the speech, language, and literacy abilities of children from infancy to adolescence. She is a certified speech-language pathologist and has experience in Early Intervention and school settings working with children from culturally and linguistically diverse backgrounds.

Daniel Hudock, PhD CCC-SLP, is an Associate Professor in Communication Sciences and Disorders at Idaho State University and Director of the Northwest Center for Fluency Disorders. His research interests are psycho-social-emotional aspects of fluency disorders, interprofessional collaborations with mental health professionals for the holistic assessment and treatment of fluency disorders, and the neuroscience of speech perception and production in people who stutter using high-density electroencephalography (EEG) time-course analysis.

Heather L. Ramsdell-Hudock, PhD CCC-SLP, University of Memphis, 2009, is an Associate Professor at Idaho State University. She leads the Infant Vocal Development Laboratory and her research and teaching interests span infant vocal development, caregiver-infant reciprocity, phonetics, phonology, the scholarship of teaching and learning, and mentoring graduate and undergraduate students through the scientific process.