Title	Assessment and treatment of short-term and working memory impairments in stroke aphasia - a practical tutorial
Author(s)	Salis, Christos; Kelly, Helen; Code, Chris
Publication date	2015-06-30
Original citation	Salis, C., Kelly, H. and Code, C. (2015) 'Assessment and treatment of short-term and working memory impairments in stroke aphasia - a practical tutorial', International Journal of Language and Communication Disorders, 50, pp. 721-736. doi:10.1111/1460-6984.12172
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://dx.doi.org/10.1111/1460-6984.12172 Access to the full text of the published version may require a subscription.
Rights	**© 2015, Royal College of Speech and Language Therapists. Published by John Wiley & Sons, Inc. This is the peer reviewed version of the following article: Salis, C., Kelly, H. and Code, C. (2015) 'Assessment and treatment of short-term and working memory impairments in stroke aphasia - a practical tutorial', International Journal of Language and Communication Disorders, 50, pp. 721-736. doi:10.1111/1460-6984.12172, which has been published in final form at https://doi.org/10.1111/1460-6984.12172 This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.**
Item downloaded from	http://hdl.handle.net/10468/7228

Assessment and treatment of short-term and working memory impairments in stroke aphasia: A practical tutorial

Running head: Memory and aphasia

Keywords: stroke, aphasia, short-term memory, working memory, assessment, treatment

Christos Salis

Newcastle University

Speech & Language Sciences

King George VI Building

Queen Victoria Road

Newcastle upon Tyne, NE1 7RU, UK

Tel: +44 (0)191 208 8875

E-mail: christos.salis@ncl.ac.uk

Helen Kelly

University College Cork

Department of Speech & Hearing Sciences

School of Clinical Therapies

Brookfield Health Sciences Complex

College Road

Cork, Republic of Ireland

Tel: + 353 (0) 21 490 1746

Email: helen.kelly@ucc.ie

Chris Code

University of Exeter

Department of Psychology

Washington Singer Laboratories

College of Life and Environmental Sciences

Exeter, EX4 4QG, UK

Tel: +44 (0) 1392 724693

E-mail: C.F.S.Code@exeter.ac.uk

Declaration of interest: None

Acknowledgements: We would like to thank the two reviewers who made several helpful, insightful and fair suggestions to previous versions of this paper. We would also like to thank Jen Dodds Vigouroux for sharing her clinical insights on an earlier version of this paper.

Abstract

Background: Aphasia following stroke refers to impairments that affect the comprehension and expression of spoken and/or written language and co-occurring cognitive deficits are common. In this paper, we focus on short-term and working memory impairments that impact on the ability to retain and manipulate auditory-verbal information. Evidence from diverse paradigms (large group studies, case-studies) report close links between short-term/working memory and language functioning in aphasia. This evidence leads to the hypothesis that treating such memory impairments would improve language functioning. This link has only recently been acknowledged in aphasia treatment but has not been embraced widely by clinicians.

Aims: To examine the association between language, and short-term and working memory impairments in aphasia. To describe practical ways of assessing short-term and working memory functioning that could be used in clinical practice. To discuss and critically appraise treatments of short-term and working memory reported in the literature.

Methods: Taking a translational research approach, this paper provides clinicians with current evidence from the literature and practical information on how to assess and treat short-term and working memory impairments in people with aphasia. Published treatments of short-term and/or working memory in post-stroke aphasia are discussed through a narrative review.

Main Contributions: This paper provides the following: A theoretical rationale for adopting short-term and working memory treatments in aphasia; It highlights issues in differentially diagnosing between short-term, working memory disorders and other concomitant impairments, for example, apraxia of speech. It describes short-term and working memory assessments with practical considerations for use with people with aphasia. It also offers a description of published treatments in terms of participants, treatments and outcomes. Finally, it critically appraises the current evidence-base relating to the treatment of short-term and working memory treatments.

Conclusions: The links between short-term/working memory functioning and language in aphasia are generally acknowledged. These strongly indicate the need to incorporate assessment of short-term/working memory functioning for people with aphasia. While the supportive evidence for treatment is growing and appears to highlight the benefits of including short-term/working memory in aphasia treatment, the quality of the evidence in its current state is poor. However, because of the clinical needs of people with aphasia and the prevalence of short-term/working memory impairments, incorporating related treatments through practice-based evidence is advocated.

What is already known on the subject: Historically, definitions of stroke aphasia have acknowledged the involvement of concomitant memory deficits, specifically, short-term and working memory. Despite these long-standing links, and contemporary neuropsychological and neuroimaging evidence, clinical practice has not considered seriously the relevance of short-term and working memory abilities in the assessment and treatment of aphasia.

What this paper adds: Aiming to influence clinical practice, we make a case for the adoption of protocols of short-term/working memory assessment and treatment by speech and language therapists as part of the rehabilitation of post-stroke aphasia through practice-based evidence. We describe practical ways of assessing short-term/working memory functioning and give a critical overview of treatment methods that have been published to date.

Introduction

That stroke often causes chronic language deficits (i.e., aphasia) resulting in communication disabilities in the spoken and written modalities are well-known. Stroke also results in memory impairments (Snaphaan & de Leeuw, 2007). The Stroke Association Stroke Survivors' Needs Survey (McKevit et al., 2011) found that 42.8% of respondents reported memory problems. The majority of respondents (59%) regarded getting help with memory problems an unmet long-term need. Although the survey did not focus on the nature of memory problems, a recent systematic review (Snaphaan & de Leeuw, 2007) found that memory problems in stroke are often verbal memory problems. In this paper we focus on verbal short-term and working memory impairments (STM and WM respectively) because of the close link between these impairments and aphasia.

The primary objective of this paper is to provide speech and language therapists with a theoretical rationale and practical information on how to assess and treat STM and WM impairments in aphasia, supported by a critical review of the relevant literature. We will show that STM/WM impairments co-occur with aphasia and, in many cases, contribute majorly to several language functions (e.g., lexical processing, sentence comprehension, reading). The review will also reveal that, in some people a STM/WM impairment impacting upon everyday communication may be present in mild aphasia or even in the absence of aphasia. Such patterns of impairment highlight the importance of assessing and treating STM and WM impairments, which we advocate together with others (Murray, 2012; Martin & Reilly, 2012). Clinical guidelines of the Royal College of Speech and Language Therapists (2005) state that components of cognition, including memory, should be considered when assessing aphasia and such impairments must be considered in all aspects of clinical management. More recently, the National Clinical Guidelines for Stroke (Intercollegiate Stroke Working Party, 2012) not only recommended the assessment of cognitive impairments in stroke, but also highlighted the speech and language therapist's role in interpreting cognitive tests in people with aphasia so as to tease apart the contribution of language impairments in relation to other cognitive impairments.

The main tenet of evidence-based practice is that clinicians only adopt the best available evidence in their practice (Greenlagh, 2014). It will become apparent that the quality of the small evidence-base for effective treatments of STM and WM impairments in stroke aphasia is very limited. With evidence-base clinical decision-making in mind, it would be premature to adopt STM and WM treatments in clinical practice. However, there is a clear need to gather evidence in order to enrich and improve the evidence-base of current treatments (Royal College of Speech & Language Therapists, 2005; Brady, Kelly, Godwin & Enderby, 2012). Dobinson and Wren (2013) make a case for practice-based evidence. This is defined as the gathering of evidence during the course of everyday clinical activity. They go on to say that a key aspect of practice-based evidence is for clinicians to know what support and resources they need and where to find them. The current paper has this objective. We also know that clinicians are often confronted with cases where there is either a suspicion, or more overt signs that a person's aphasia involves impaired STM and WM, which either interfere with everyday functioning or, equally important, with treatments that the clinician delivers. In short, we make a case for practice-base evidence in order to improve the quality of future treatments of STM/WM. This avenue of clinical enquiry is likely to lead to improved patient care and address the needs of stroke survivors (cf., McKevitt et al., 2011).

The next section explores the relationship between STM, WM and related concepts. In following sections, we show that STM and WM impairments are integral to aphasia, thereby, providing a theoretical rationale through a historical lens. This is followed by a description of measures of STM and WM functioning, based on relevant studies. Finally, we present the findings of a literature review focusing on description of participants, treatment methods and outcomes.

STM and WM: Definitions, evolution of concepts and terminological differences

Verbal STM and WM are closely related and interface with other cognitive functions, notably, attention. In a historical overview of STM and WM, Cowan (2008) states that at the end of the 19th century a conceptual distinction, originating with William James (1890) but still relevant today, was drawn between primary and secondary memory. Primary memory referred to the ability to remember information received through auditory and visual channels for a brief period of time, immediately after this information was registered. In contrast, secondary memory characterised the ability to remember over a much longer time period. In the 1950s and 1960s, the terms primary and secondary memory were gradually replaced by short- and long-term memory respectively. There are two main differences between STM and long-term memory: First, information in STM is short-lived and can last up to 30 seconds, a feature known as temporal decay. Second, STM has limited capacity in terms of encoded information memory units (i.e., chunks) (Cowan, 2008). WM refers to a temporary memory system used for mentally manipulating information (Baddeley, 2012). STM is also a temporary memory system that governs the ability to recall information immediately after its presentation in a relatively unprocessed state, that is, without mental manipulation (Baddeley, 2012). So, WM entails mental manipulation of information while STM is about recalling information without such manipulation. Verbal STM is usually tested with verbal items (e.g., digits) presented in a list format. Non-verbal STM is assessed using non-verbal items (e.g., visuo-spatial), with tasks utilising three-dimensional blocks such as the Corsi block task. Assessment can also involve non-verbal sounds. Recall is usually carried out in a serial manner, from the beginning to the end of a list (Baddeley, 2012). However, recall could also be in any order (i.e., free recall). Serial recall assesses STM for order information while free recall assesses STM for item information. Later on, we discuss these two different recall methods in more detail. The rehearsal process whereby one repeats sub-vocally using inner speech, in one's head as it were, a sequence of words from beginning to end does not qualify for mental manipulation. However, if this sequence of words needs to be sorted from the end of the sequence to the beginning (backward recall), this entails mental manipulation requiring WM.

The well-known multi-component model of WM (shown in figure 1) comprises several stores or buffers that are responsible for auditory-verbal and visuo-spatial information over the short-term, as well as an episodic buffer (Baddeley, 2012). There is also an attentional mechanism, known as the 'central executive', which utilizes information that is temporarily held in a number of stores, or 'slave systems' as they became known, including a verbal 'phonological loop', and a visual 'visuo-spatial scratchpad' or 'sketchpad'. Both slave systems and the central executive have limited capacity – the slaves only hold a finite amount of information; the central executive is limited in terms of the amount of information it can process at one time and the speed with which it processes the information (Baddeley, 2012). Again, precise capacity in terms of quantity of information as well as retention time varies, but 4 seconds is often considered the upper limit, or typically 2 to 6 items, depending on the verbal material and the task, although there are large individual differences in these metrics in healthy individuals (Cowan, 2008). The phonological loop was originally known as the ‘articulatory loop’ because it processes spoken information. Finally, the ‘episodic buffer’ was added in order to link STM and long-term memory. The episodic buffer holds integrated episodes in a multidimensional space, linking components of WM and also linking WM to perception and long-term memory (Baddeley, 2012).

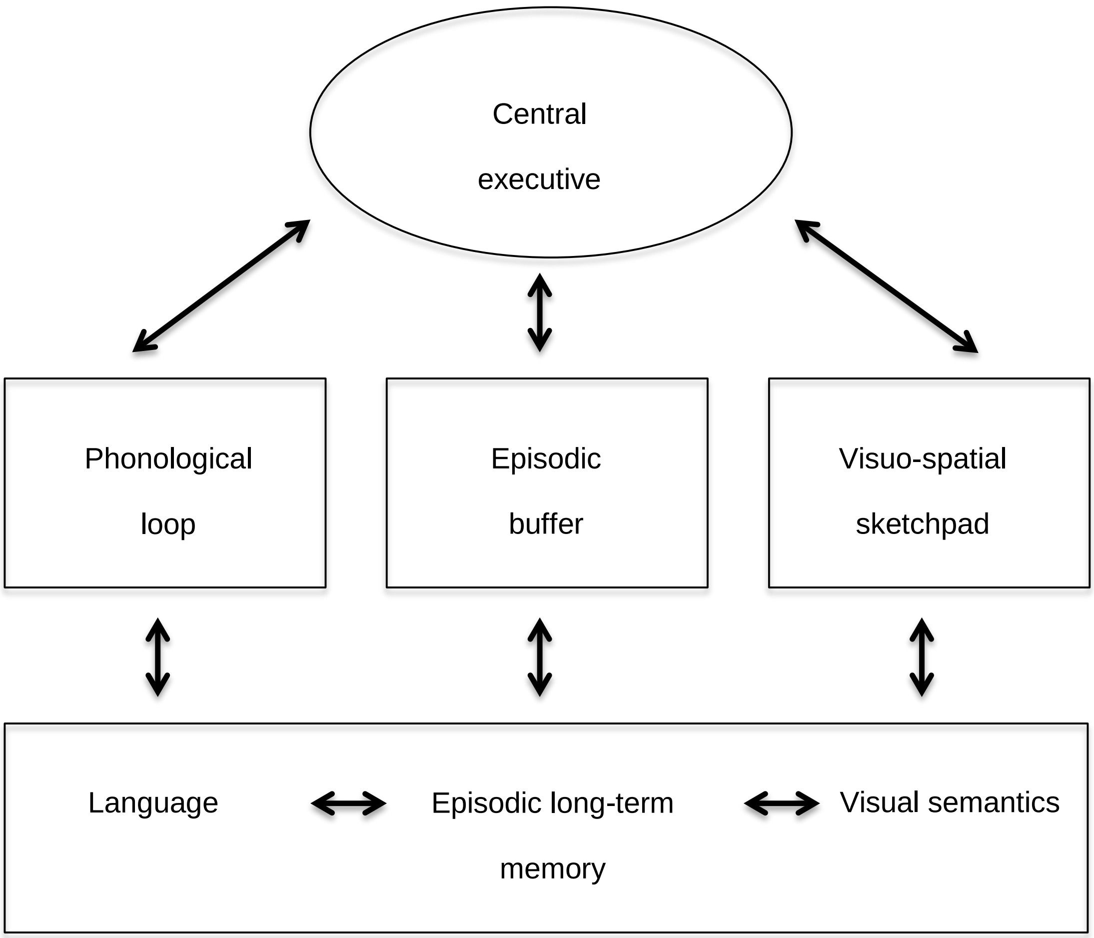

Figure 1. The multi-component model of working memory

While the role of attention and the link between STM and long-term memory are part of Baddeley's multi-component model, in other theories of WM attention plays a more central and explicit role. In what is known as the 'embedded processes model' (Cowan, 2008), WM is viewed as the activated portions of long-term memory that fall within the focus of attention for a particular task. STM is derived from a temporarily activated subset of information in long-term memory. The amount of activated information is limited (about 4 chunks) and decays quickly (in about 4 seconds). The activation is achieved by attention control processes and the activated subset decays if not rehearsed.

Inhibition is a more specific executive-attentional mechanism, which is argued to play a key role in the ability to encode, store and retrieve information (e.g., Hasher, Lustig, & Zacks, 2007). Inhibition refers to the ability to resist interference or exert voluntary inhibitory control over information that is not relevant for a particular task. Inhibition in language processing reflects our ability to consciously resist unwanted language activity. Inhibition is essential for resisting distraction, sharpening the focus of attention for a given task and, ultimately, ensuring that information remains active and accessible in WM (Hasher et al., 2007). Efficient inhibitory control filters out irrelevant stimuli and helps focus attention on task-relevant ones. For example, when we try to focus attention on reading the text on a webpage, crowded with static pictures, flashing images, mixed font sizes and colours of text, we have to inhibit all distracting information in order to maximise meaning extraction from the written text.

What emerges from these different perspectives is that STM and WM are related processes and are also related to long-term memory in that they support the creation of long-term storage of memories and also facilitate access from long-term memory to a readily available state. Attention control mechanisms, conceived as elements of the central executive entailing focus of attention and inhibition, are key in WM efficiency across theories of WM. However, controversy surrounds STM and WM, especially, in the nature of the relationships between the central executive component and the cognitive processes that operate therein (notably, attention and executive functioning). There is on-going work and debate is attempting to refine our understanding of the nature of the relationship between these processes and speech and language.

The close link between aphasia and STM/WM

The ‘standard’ theoretical and clinical (as opposed to the social) model sees aphasia primarily as a set of inter-related linguistic impairments often affecting multiple levels of linguistic description such as phonology, semantics, morphology, syntax, in spoken and written modalities (e.g., Whitworth, Webster & Howard, 2013). Although this model has been useful, a conceptualisation of aphasia based on linguistic description alone helps understand only part of the impairments. In this section, we outline how STM and WM impairments interact with language and discuss related diagnostic implications.

Historically, aphasia was commonly seen as a memory disorder up until, and including Broca, and memory impairment has regularly been invoked in a variety of aphasic symptoms. The Greco-Roman physician Galen (130-200 AD) described the three ventricles as the home of human faculties, based on the even earlier ideas of Herophilos (332-280 BC), 'the father of anatomy', and by the middle ages this idea dominated as 'cell theory', which believed aphasia resulted from damaged memory processing in the third ventricle, or cell (Finger, 1994). Aphasia as an impairment of some aspect of memory continued as an explanation well into the 18th century: Johann Gesner (1788-1801) described what many consider the earliest complete theory of aphasia in his paper 'The Language Amnesia' (Tesak & Code, 2008). Jean-Baptiste Bouilaud (1796-1881), the supporter of Franz Josef Gall's localisation theory, and the dominant physician in Paris during Broca's time, divided his cases into those with articulation disorders and those with language disorders, the basis of which was an impairment of memory. In the 19th century, Paul Broca's own classification of aphasia included 'verbal amnesia' where patients had forgotten the meaning of words, which would later be termed 'sensory aphasia' by Bastian and Wernicke (Tesak & Code, 2008). In Wernicke's original sensory-motor model of language, words are stored as two types of memory 'images': Motor-movement and sound-memory images, which became the core components of the 19th century model of aphasia. We recall too that Bateman, Luria, Eisensen and others had 'amnesic' or 'amnestic aphasia' in their classifications, what is commonly referred to in contemporary times as anomia.

Modern studies highlighted the strong connection between language and STM impairments (Schuell, Jenkins & Jimenez-Pabon, 1964; Crocket, Clark, Spreen &

Klonoff, 1981; Beeson, Bayles, Rubens & Kaszniak, 1993; Leff et al., 2009; Laures-Gore, Shisler-Marshall & Verner, 2011). Schuell et al. (1964) and Crocket et al. (1981) found that digit repetition (a common test of STM, also known as forward digit span) was a key predictor of language performance (receptive and expressive). More recently, Leff and colleagues (2009) reported that digit repetition predicted performance in spoken sentence comprehension and argued for the left posterior superior temporal gyrus and sulcus as mediating auditory STM and sentence comprehension. Correlations between WM measures and aphasia severity have also been reported (Caspari, Parkinson, LaPointe, & Katz, 1998; Potagas, Kasselimis, & Evdokimidis, 2011). For example, Caspari and colleagues (1998) reported moderate to strong correlations between WM measures (the reading and listening span tasks, described later) and performance on the Western Aphasia Battery (Kertesz, 1982) as well as the Reading Comprehension Battery for Aphasia (LaPointe & Horner, 1978), showing that WM underpins spoken and reading skills in aphasia. Caspari et al. did not include a control group. This means that it cannot be determined if the aphasic participants' performance was indeed deficient, although it is likely to have been. People with aphasia perform poorly in tests of STM and WM, which may have led Schuell and colleagues (1964) to define aphasia as a dual impairment characterised by anomia and impaired verbal retention span (i.e., STM impairment). In all these studies STM and WM measures required speech output. This could potentially confound the measurement of STM and WM because of possible articulatory deficits and anomia. STM tasks that do not require speech output such as pointing span (described later) also revealed STM deficits across classical aphasia syndromes (Goodglass, Berko Gleason, & Hyde, 1970; Martin & Saffran, 1999; Martin & Ayala, 2004). It is clear then, that if STM and WM measures are used in the assessment of aphasia, STM and WM impairments emerge and are a common feature in most people.

In the previous section, we introduced the concept of the memory ‘buffer’ in relation to the multi-component model of WM (figure 1). The buffer is not specific to that model but appears in several theoretical frameworks of normal as well as impaired speech and language processing. It refers to a STM mechanism that stores information (serially or not) from a particular modality, spoken or written, input or output. The buffer has been studied many times, and, indeed, separate buffers have been invoked to operate in a range of language tasks - articulatory buffer, graphemic buffer, lexical buffer (Whitworth et al., 2013). One view of apraxia of speech, for instance, is that the halting speech observed may result from a reduction in articulatory buffer capacity so that people with the condition are only able to programme a smaller chunk of speech, as little as a single syllable (Rogers & Storkel, 1999). Martin, Lesch, and Bartha (1999) have shown that semantic information contributes to STM and that input and output phonological codes are maintained in separate buffers. In naming, shared pathways that underpin retrieval of phonology from semantics, also underpin feedback from semantics in STM tasks (Martin et al., 1999). The buffer is also invoked in written language production. In the version of the cognitive neuropsychological characterisation of aphasia, there is a graphemic output buffer (Whitworth et al., 2013), responsible for maintaining graphemic representations before they are realised in manual-motoric patterns. The distinctive feature of buffer impairments, not only regarding the graphemic, but also the phonological output buffer, is length effects, whereby longer words (written and spoken respectively) would be more prone to errors than shorter words (Whitworth et al., 2013).

The questions arise as to what the essential relationship between language and memory (STM/WM) processing in cognition and aphasia actually is and why it is that we commonly see the co-occurrence of language and STM/WM impairment in aphasia? These issues were examined in aphasia studies utilizing Dell's (1986) connectionist model (Martin, Saffran & Dell, 1996; Dell, Schwartz, Martin, Saffran & Gagnon, 1997). The fundamental elements of this model are connection strength, in terms of weights on units and decay rate. Martin et al. (1996) developed a connectionist model that successfully simulated changes in error type in their participant NC's recovery of naming and repetition, through a process of 'lesioning'. Lesioning caused changes to decay rate. Improvements (i.e., decreases) in decay rate with recovery appeared to underlie recovery of repetition and naming. As decay rate decreased in naming, there was a large decrease in formal (sound-related) paraphasias and a smaller decrease in semantic paraphasias, reflecting less use of phonological specifications and easier access to semantics. However, as decay rate decreased in a repetition task, there was a reduction in semantic paraphasias and an increase in formal paraphasias as NC became less dependent on semantics. Dell et al. (1997) examined the deficits in a case-series of people with aphasia (N=21) on a picture-naming task. They were able to simulate errors in picture naming fairly accurately. Analysis revealed that the model fitted each participant's deficits and led them to a classification scheme where each participant's deficits could be characterised in terms of the two key parameters of connection strength and decay rate. Martin and Gupta (2004, p. 14) make the point that 'on this proposed severity continuum, it is the case that all individuals with aphasia should also present with verbal STM deficits, but not all individuals with verbal STM deficits should present with obvious aphasia'.

To summarise, throughout the history of aphasia, STM has not only been an explanatory mechanism for a range of language abilities, but also a function closely linked to language disability following brain damage, which in many cases is a predictive factor. As Martin and Reilly (2012) note, the content of verbal STM/WM memory is language.

Impairments of STM and WM in aphasia

In general, impairments of STM and/or WM in aphasia can be defined and also identified by below age- and education-appropriate performance in a STM and/or WM test, provided the person with aphasia has understood the demands of the test and is able to cope with the speech or other test demands as we discuss in the next section. It is possible that a severe STM impairment can be present in people with mild aphasia (Warrington & Shallice, 1969; Martin & Gupta, 2004) and that STM and WM impairments can be discerned even when speech demands are minimal (e.g., Martin & Ayala, 2004). In this section, we focus on two distinct STM impairments (phonological and semantic) that can occur in aphasia. In as far as lexical processing can break down along phonological and semantic dimensions in aphasia (Whitworth et al., 2013), STM can also break down along similar dimensions (Martin & Ayala, 2004; Howard & Nickels, 2005; Martin & Allen, 2008).

Phonological and semantic STM impairments describe impairments that relate performance in phonological and semantic tasks to the way these underlying codes influence different aspects of STM (phonological and semantic respectively). We discuss three approaches that attempt to diagnose such difficulties. While all approaches share similarities in that phonology and semantics are implicated in STM performance, the diagnostic process is based on different assessment tasks.

The first approach comes from N. Martin's group. Martin and Ayala (2004) showed that phonological and semantic skills for lexical processing impact in a different way upon STM. They divided a group of 47 speakers with aphasia into two subgroups based on their performance on standard phonological (e.g., phoneme discrimination, spoken rhyme judgements) and semantic processing (e.g., synonym judgements) tasks. The speakers were then tested on verbal (e.g., digit span, pointing span) and non-verbal (e.g., Corsi block) STM measures - we describe the verbal STM measures in more detail in the next section. Some participants repeated in serial order while others did not. There were two potentially diagnostic patterns of performance relating to whether participants were unable to recall items at the beginning of the list (absence of primacy effect) and whether participants were unable to recall items at the end of the list (absence of recency effect). Lack of primacy effect was associated with poor access to lexical semantics and/or reduced semantic STM, whereas a lack of recency effect suggested that access to phonology or phonological STM is impaired.

Similar diagnostic distinctions between phonological and STM impairments were presented by Howard and Nickels (2005). Participant HB was better at recalling lists of high imageability than low imageability words. HB did not show a word length effect (a diagnostic feature of phonological STM impairment) and was able to repeat lists with one- and three-syllable words equally well. Participant MMG showed an effect of word imageability (a diagnostic feature of semantic STM impairment) for three-word lists but not with two-word lists. Another difference in the two participants emerged when they were presented with spoken and written letter lists. Phonologically similar lists were drawn from rhyming (e.g., P, V, B, T) and non-rhyming letters (e.g., J, S, H, Y). HB did not show a phonological similarity effect with either auditory or visual presentation (another diagnostic feature of phonological STM impairment).

The third approach comes from R. Martin's group. The diagnostic features of phonological and semantic STM impairments (R. Martin & Allen, 2008) are shown in table 1.

Table 1. Diagnostic features of phonological and semantic STM impairments

Phonological STM impairment	**Semantic STM impairment**
• Good single word processing	• Good single word processing
• STM span of 1 to 3 items	• STM span of 1 to 3 items
• Better performance in semantic probe than rhyme probe tasks	• Better performance in rhyme probe than semantic probe tasks
• Better performance in span tasks with short than long words (i.e., number of syllables)	• Similar performance in span tasks with short and long words (i.e., number of syllables)
• Better performance in span tasks with words than non-words	• Similar performance between word and non-word span tasks

• Better performance when information is presented in written than spoken modality	• Better performance when information is presented in spoken than written modality

An important feature is differentiable performance in phonological and semantic STM tasks (R. Martin & Allen, 2008) (table 2). In a phonological STM task, the person listens to lists of words. At the end of each list, the person listens to an additional probe word. In the phonological condition s/he has to decide if a probe word rhymes with any of the words presented in the list. In the semantic probe task, the person listens to lists of words. At the end of the list, s/he has to decide if a probe word belonged to the same semantic category as any of the words in the list.

Table 2. Examples of phonological and semantic probe tasks (three item lists)

	Word lists	**Probes**
Phonological	chair, box, shoe	Does the word ‘hair’ rhyme with any of the words you have heard?
Semantic	pen, train, bird	Is the word ‘car’ in the same category as any of the words you have heard?

To date, the differential diagnosis of phonological and semantic impairments has been utilised explicitly in a treatment study by Harris, Olson and Humphreys (2014). We should point out that the term ‘phonological’ as used by these and often other authors refers to processing of non-words and the term semantic refers to processing of real words. Some of the knowledge about phonological and semantic STM impairments comes from studies following the single-case study paradigm that partially defines cognitive neuropsychology. While often rare in clinical practice to come across people with aphasia who show differentiable and often striking patterns of language processing deficits, the single-case study paradigm provides the framework for the building blocks of verbal STM architecture and how it relates to other aspects of language processing.

Assessing STM and WM in clinical practice

Typically, verbal STM tasks (known as span tasks) assess the ability to recall auditory-verbal (or visual) information from lists of words (e.g., digits, letters, other nouns), usually in a serial order (immediate serial recall). There are also other types of STM tasks where recall can be in any order (i.e., free) rather than serial. Free recall tasks assess STM for item information as opposed to order information (serial recall). WM tasks, inherently more complex, involve the ability to temporarily store while mentally manipulating verbal information. Before we go on to describe some of these tasks, we consider the potential impact of other impairments, common after a stroke, which may confound assessment of STM/WM.

Most STM and WM tasks rely on speech repetition, something that many people with aphasia have difficulties with, even at single word level. The extent to which production difficulties confound STM and WM has not been studied systematically. It is worth bearing this in mind when using memory tasks that require speech output. In a review of memory impairments following stroke, Snaphaan and de Leeuw (2007) noted the exclusion of people with severe strokes and people with aphasia from memory assessments. These groups could not complete testing with digit repetition, the most commonly used task in the reviewed studies. The presence of motor speech impairments may make the differential diagnosis between aphasia, STM/WM and motor speech impairments difficult. Also, some tasks, especially WM tasks, require understanding of complex verbal instructions making them unsuitable for people with lower levels of comprehension. The presence of sensory and/or perceptual visual deficits such as neglect in visual tasks (e.g., pointing span) may interfere with performance. We are aware that most speech and language therapists work in multi-disciplinary teams, so some of the assessments we describe may already be carried out by other professionals, such as occupational therapists and clinical psychologists. Therefore, the need for sharing relevant assessment results between professionals is clear, even if the disciplines are different. In the descriptions below we restrict ourselves to verbal STM/WM tasks. Several studies have used non-verbal/non-linguistic STM and WM tests in the aphasia literature (e.g., Gordon, 1983; Mayer & Murray, 2012).

Short-term memory tasks

Forward recall tasks: In forward recall, the person hears a list of words (usually digits or letters), which s/he then repeats. The lists increase in length and consequent difficulty as the test progresses. Forward recall can also be tested without the need for spoken production but by pointing (described in the next paragraph). In forward recall, lexical items are presented at a rate of one word per second, each item in a distinct intonation unit to prevent chunking as much as possible. Immediately after presentation, the person has to either repeat (or point to) all list items serially from the beginning to the end of the list. The following published test batteries contain forward digit recall tasks that require speech: Comprehensive Aphasia Test (CAT, Swinburn, Porter & Howard, 2005), Wechsler Adult Intelligence Scale IV (Wechsler, 2010), Wechsler Memory Scale IV (Wechsler, 2010), and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS, Randolph, 2012). The test of Memory and Learning (TOMAL-2, Reynolds & Voress, 2007) assesses forward recall span with digits as well as letters.

Forward recall tasks that require pointing are sometimes called ‘pointing span’ (Goodglass et al., 1970; Martin & Saffran, 1999; Martin & Ayala, 2004). DeDe, Knilans, Ricca and Trubl (2014) present norms from younger (n=24) and older (n=24) healthy adults. The person listens to a list of words, while looking at a blank sheet of paper. Immediately after the spoken presentation, 9 pictures appear on a different page depicting the presented words in random order. The person has to point to the pictures in the same order as the spoken words (forward recall). The same words are used throughout the test but the order of pictures differs from trial to trial. In the version standardised by DeDe et al., if the person cannot recall the order of a word they have to point to the word “blank”, written on a separate sheet of paper. It is unclear if a vocabulary check has to be carried out first. In another version of the pointing span (Goodglass, et al., 1970), a vocabulary check is first carried out to eliminate confounds of word comprehension deficits.

Underlying processes tapped by forward recall tasks: Using digits and nouns, Martin and Ayala (2004) compared performance between repetition and pointing span tasks in people with aphasia representing a spectrum of profiles but without motor speech impairments. Visuo-spatial STM was also assessed with the Corsi block. The verbal STM task was a serial order repetition task. Participants were asked to repeat items and point to items in serial order. The main findings were as follows: Similar performance between repetition and pointing span; span for digits was greater than span for words in both repetition and pointing; repetition span correlated with measures of phonological processing, while the pointing span correlated with both phonological and lexical-semantic measures; finally, visuo-spatial STM correlated with performance in repetition and pointing, suggesting the use of verbal encoding in remembering visuo-spatial information. It is worth pointing out that factors that affect processing of single words (e.g., frequency, lexicality) also affect STM capacity, a pattern that shows the close relation between word processing, STM and long-term memory (Stuart & Hulme, 2009).

Another forward recall task is sentence repetition. As we will see in the next section, sentence repetition is one of the most popular treatment tasks in STM treatments. Sentence repetition is part of several clinical tests (e.g., Comprehensive Aphasia Test, Swinburn et al., 2005; Western Aphasia Battery, Kertesz, 2006; Boston Diagnostic Aphasia Examination, Goodglass, Kaplan & Barresi, 2001). Sentence repetition can be considered a forward recall task in which words are repeated serially. Naturally, then, it can be used to assess STM for sentences, for example, by contrasting performance between 3 and 8 word sentences. Other things being equal, if a person's performance is worse with 8 words than 4 words, this could be attributed to a STM impairment. However, words in sentences are not arranged randomly but in a hierarchical structure, constrained by syntactic and semantic rules. Also, sentences contain different word classes, notably, content and function words. Speech motor demands are also relatively high in sentence repetition. These factors should be considered when interpreting performance in sentence repetition tasks, and when one compares performance between word and sentence repetition tasks. The sentence repetition subtest of the CAT (Swinburn et al., 2005) takes into account syntactic complexity to some extent.

The Token Test: This standardised test, originally developed by De Renzi and Vignolo (1962), has several shortened versions, including a popular clinical version by McNeil and Prescott (1978). Although the Token Test is not a STM test per se, but a sentence comprehension test, the hierarchical nature of the stimuli is well-suited to the assessment of STM in aphasia (Francis, Clark, & Humphreys, 2003; Salis, 2012). It involves listening to and carrying out spoken commands of increasing length and, in some subtests, grammatical complexity. The increasing length of commands in the relative absence of confounds such as vocabulary and sentence structure in subtests 1 to 4, makes this test suitable for STM assessment for item information as opposed to order. The person looks at an array of different object tokens, the main vocabulary, identified by size (*small, big*), shape (*square, circle, triangle*) and colour (*red, green, blue, black, white*). The test is divided into subtests of increasing difficulty (table 3). For subtests 1 to 4, the vocabulary and sentence structure are simple.

Table 3. Examples of the Token Test (McNeil & Prescott, 1978)

Subtests	**Examples**	**Visual demands**
1	touch the black circle	10 tokens
2	touch the big green square	20 tokens
3	touch the big green square and the little black square	10 tokens

4	touch the big green square and the little black square	20 tokens

Better performance in subtest 1 and poorer performance in subtest 3 could be attributed to a STM impairment, provided the person understands the individual vocabulary items (Francis et al., 2003). Similarly, better performance in subtest 2 and poorer performance in subtest 4 could be attributed to a STM impairment. It is important to note that subtest 1 and 3 use fewer tokens (10 as opposed to 20) than subtests 2 and 4, and consequently, impose fewer visuo-spatial demands, which may affect STM. Lesser (1976) showed that the sensitivity of the Token Test appears to have much to do with STM. Most of the test commands have information content of six items, requiring significant STM skill. While the test is able to detect aphasia, in particular, spoken comprehension impairments, it is probable that it detects STM impairments, which may contribute to the understanding of sentences with more complex and varied grammatical structures. This would warrant further testing with a spoken comprehension of grammar test.

Working memory tasks

Compared to STM tasks, WM tasks are more complex in terms of what the person is asked to do, and involve considerably greater cognitive demands. Serial recall is not a feature of WM tasks as it is in STM tasks. Another feature of certain WM tasks is that of ‘updating’. Updating refers to the process of revisiting previously encoded/stored stimuli and is an attentional-executive component. Tasks such as the n-back, digit sequencing and alphabet spans, described below, have an updating component.

Backward recall: Unlike forward recall, backward recall involves the presentation of lists of spoken words, usually digits or letters that increase in length, which need to be recalled from the end of the list to the beginning (i.e., mental manipulation demand). Backward repetition can be found in published batteries like the Wechsler Adult Intelligence Scale IV (Wechsler, 2010) and Wechsler Memory Scale IV (Wechsler, 2010) and RBANS (Randolph, 2012). Backward recall can also be carried out using pointing.

Listening span: The so-called listening span task focuses on processing and storage of verbal information in WM. Tompkins and colleagues (Tompkins, Bloise, Timko & Baumgaertner, 1994; Lehman & Tompkins, 1998) adapted this task and used it with people with aphasia. In this task, the person listens and responds to sets of simple statements (e.g., *chickens lay eggs*, *trains can fly*). After each statement, the person states if it was true or false, forcing understanding of the statement (processing component). At the beginning, the person is instructed to retain the last word of the statement (storage component) because s/he would have to verbally recall the last word of each statement in any order at the end of the set. So, the correct answer for the example would be *eggs, fly* (two item set). The sets gradually increase in number of statements and consequently in number of words that have to be recalled at the end of each set. The scoring method involved addition of errors from both recall and true/false verification of sentences. Lehman and Tompkins (1998) provide the actual test items. Normative data were elicited from 28 healthy older adults (mean age 62 years).

Digit/word span sequencing: This task is part of Wechsler Adult Intelligence Scale IV (Wechsler, 2010) and Wechsler Memory Scale IV (Wechsler, 2010). In this task, a series of digits is spoken. The person has to reorder the digits in ascending order numerically. Digit span sequencing is a new feature of the Wechsler test batteries and was added to increase the WM demands of the digit span subtest (forward and backward). A test similar in principle is the List Sorting Test which involves sorting noun referents in size order (e.g., elephant, bird, ant; Tulsky et al., 2014).

N-back: In this task, people hear words (stimuli could also be sentences) or see pictures, one by one, presented as a continuous list. The person needs to keep track of the presented stimuli and retain the stimuli that s/he heard (or saw) at a pre-specified interval, usually 1, 2 or 3 stimuli (this is referred to as *n*). The person needs to monitor and respond when a stimulus appears 1, 2 or more stimuli (n) earlier (back). Table 4 shows an example of 1-back. Different versions of the n-back task were used to investigate WM performance of people with aphasia by Christensen and Wright (2010), Mayer and Murray (2012) as well as DeDe et al. (2014). DeDe and colleagues (2014) provide normative data (younger and older healthy adults) from 1-back and 2-back task. Test-retest reliability was .86 in the 1-back task but .35 in the 2-back task, suggesting that the 2-back task is not a stable measure. With regards to n-back, Mayer and Murray (2012) comment as follows: The flexibility in terms of stimuli (verbal, non-verbal), response type, and inter-stimulus intervals are also one of the greatest weaknesses of n-back. This is because it is difficult to assume that the feasibility and reliability estimates for one version of the n-back will be consistent across other versions of the task. This statement was corroborated by by Christensen and Wright (2010). In Mayer and Murray's study, the n-back task indicated good clinical feasibility and high estimated test-retest reliability for their n-back task as utilized for adults with aphasia, consistent with reliability data of the n-back for healthy speakers. However, the data from DeDe et al. (2014) suggest otherwise.

Table 4. Example of a 1-back task

Person hears	3	1	4	7	2	9	4
Correct response	None	None	3	1	4	7	2

Gold standard WM tasks: DeDe et al. (2014) describe three WM tasks (alphabet span, subtract-2 span and reading span) and calculated a composite score across these three tasks, which, they termed as "gold standard". The choice of tasks was influenced by Waters and Caplan (2003) who showed that, in healthy adults (age range 18-80 years), test-retest reliability improves when scores are combined. Test-retest reliability coefficients were higher than .80 in each of the three tasks. The coefficient for the composite score was .92. Brief descriptions of the tasks are provided below.

In the alphabet span, sets of monosyllabic words are verbally presented by the clinician. The lists become progressively longer (2 to 8 words), although it is not clear how many trials per list length are presented. The person has to rearrange the words alphabetically. For example, “pen”, “mug”, would be arranged as “mug”, “pen”.

In the subtract-2 span, lists of digits are spoken by the clinician at a rate of one digit per second. The person has to repeat the digits in the same order by subtracting 2 (hence minus 2) from each spoken number. So, the correct answer for the digits 9-3-5 would be 7-1-3. The lists begin with two digits and continue to lists of eight digits. It is not clear how many trials per list length are presented.

The reading span task is similar to the listening span task (Lehman & Tompkins, 1998) described earlier. The difference is that the sentences are written. The person has to read and verify each sentence (that is, decide if the sentence makes sense or not), and recall verbally the last word of each sentence (presumably in any order). In this version, the shortest set contains two sentences (span 2) and the longest six sentences (span 6).

In this section, we described STM and WM tasks that could be used to assess STM and WM functioning in aphasia. Tasks that do not require speech output (e.g. pointing span, n-back) could be used to assess STM and WM in people who have not only aphasia but also motor speech impairments. The tasks we described have in our view relatively good psychometric properties in terms of test-rest reliability. Test-retest reliability is particularly important because a high reliability coefficient (greater than .75) suggests that the test is stable over time. Commercially published tests (CAT, Wechsler, RBANS, TOMAL-2) have fairly large standardisation samples and therefore more likely to be representative of a population. In principle, they should be preferred over other tasks. However, in some cases the normative samples have been small (Lehman & Tompkins, 1998; DeDe et al., 2014). The ultimate choice is likely to be influenced by the clinical profile of the person with aphasia and his/her communicative needs. While the search for a gold standard in STM/WM assessment has begun (DeDe et al., 2014), the field is a long way of a true gold standard for STM/WM assessment in aphasia (cf., Greenlagh, 2014).

Treatment studies: Participants, tasks and outcomes

There are two key features of STM/WM treatments that distinguish them from more established treatments: First, the use of STM and WM tasks in a treatment protocol, and, second, measurement of STM/WM skills before and after treatment. Published treatments of STM and WM in stroke aphasia reported in the literature are few (summarised in table 5). We discuss these studies in terms of the participants, treatment tasks and outcome measures used. We excluded published treatments that used STM/WM tasks but have not included pre- and post-treatment measures of STM/WM. The overarching rationale for treating STM/WM memory in all studies is that the impaired language processing skills (spoken, written) were considered to be, partly, due to impaired STM and/or WM.

Participants: To date, 14 participants whose stroke-related STM/WM impairments received specific treatment have been reported in the literature. Of those participants, 5 are reported to have recovered language functions as determined by performance within normal limits on standard aphasia tests (Mayer & Murray, 2002; Majerus, van der Kaa, Renard, van der Linden & Poncelet, 2005; Vallat, Azouvi, Hardisson, Meffert, Tessier & Pradat-Diehl, 2005; Harris, Olson & Humphreys, 2014; Vallat-Azouvi, Pradat-Diehl & Azouvi, 2014). Nonetheless, these participants self-reported language difficulties in reading longer pieces of text (Mayer & Murray, 2002; Vallat-Azouvi et al., 2014), mental arithmetic tasks (Majerus, et al., 2005; Vallat-Azouvi et al., 2014), difficulties with memory and attention (participant AK in Harris et al., 2014), difficulties in multi-talker conversations (Majerus et al., 2005; Vallat et al., 2005). The remaining participants presented with frank aphasia and STM/WM impairments. Seven participants were speakers with conduction aphasia (Vallat et al., 2005; Murray et al., 2006; Koenig-Bruhin & Studer-Eichenberger, 2007; Kalinyak-Fliszar, Kohen & Martin, 2011; Berthier et al., 2014), one was transcortical motor (Salis, 2012) and one presented with Broca's aphasia (Harris et al., 2014). The treatment outcomes are summarised in table 6.

Table 5. Summary of STM/WM treatment studies (in chronological order)

Studies	Participants	Main difficulties
Mayer & Murray (2002)	N=1, male, age=62	text comprehension
Francis et al. (2003)	N=1, female, age=69	spoken sentence comprehension
Majerus et al. (2005)	N=1, female, age=50	digit retention, mental calculation, comprehension of multi-talker discourse
Vallat et al. (2005)	N=1, male, age=53	comprehension of multi-talker discourse; note taking

Murray et al. (2006)	N=1, male, age=57	comprehension of multi-talker discourse, comprehension of complex and figurative language
Koenig-Bruhin & Studer-Eichenberger (2007)	N=1, male, age=47	word finding, discourse production
Kalinyak-Fliszar et al. (2011)	N=1, female, age=55	semantic and phonological processing, naming, spoken sentence comprehension
Salis (2012)	N=1, female, age=73	spoken words, sentence, discourse
Harris et al. (2014)	N=2, both male, ages=73 (DS), 74 (AK)	DS: naming, spoken sentence production. AK: non-word reading, episodic memory, auditory attention
Berthier et al. (2014)	N=3, all male, ages=51 (RRM), 52 (VRG), 72 (JTO)	RRM: naming, mild speech apraxia, spoken sentence comprehension. VRG: spoken sentence

		comprehension. JTO: non-word and sentence repetition
Vallat-Azouvi et al. (2014)	N=1, female, age=38	text comprehension, note taking, memorising information, conversation

Table 6. Summary of outcomes of STM/WM studies

Studies	Summary of outcomes
Mayer & Murray (2002)	Reduction in word recall errors in the sequenced exercises of WM. Post-treatment performance in a WM test was within one standard deviation of normal performance. Maintenance: Two months after the end of treatment, participant's performance was similar to performance during treatment in treatment tasks (WM and reading).
Francis et al. (2003)	Improvements in forward and backward digit span, sentence repetition (in terms of number of words remembered), recognition memory test, Token Test (in terms of sentence length and complexity).
Majerus et al. (2005)	Improvements in digit span, word span (proportion of words recalled), non-word repetition (only for high phonotactic frequency items), non-word rhyme judgements.

Vallat et al. (2005)	Improvements in forward digit span, spoken recall of spoken consonants at delay intervals of 5, 10 and 20 seconds, two ecological questionnaires: 1. Assessing the consequences of WM impairments on daily-life activities, 2. A verbal communication questionnaire. Participant was able to return to previous occupation after treatment.
Murray et al. (2006)	Improvements in accuracy and latency in attentional tasks, modest improvements in a WM test, trend towards improvement in comprehension and processing speed of paragraphs.
Koenig-Bruhin & Studer-Eichenberger (2007)	Improvements in digit span, phonological and semantic recognition tasks, matching listening span task, mean sentence length.
Kalinyak-Fliszar et al. (2011)	Improvements in word and non-word repetition, word pair repetition, sentence repetition, rhyming and synonym judgments, naming, non-word substitution errors in naming.
Salis (2012)	Improvements in forward digit span, matching listening span, Token Test, a sentence comprehension test.

Berthier et al. (2014)	For all participants (RRM, VRG, JTO) there were improvements in overall language abilities as measured by the Western Aphasia Battery, connected speech, repetition of word triplets. RRM and VRG improved in word pair repetition but JTO did not.
Harris et al. (2014)	For both DS and AK there were improvements in real and non-word recall after both treatments (phonological, semantic). In sentence anomaly judgments, DS improved after semantic treatment but not after phonological. However, AK did not improve after either treatment. In sentence repetition, AK improved but DS did not. In sentence comprehension, DS improved after semantic but not after phonological treatment. AK showed no change.
Vallat-Azouvi et al. (2014)	Improvements in forward and backward digit span (auditory and visual), recall of spoken and visual stimuli at delay intervals of 5, 10 and 20 seconds, Token Test, improvements in a questionnaire assessing everyday life problems related to working memory deficits.

Treatment tasks: The WM task used by Mayer and Murray (2002) involved reading and was similar to the listening span task (Lehman & Tompkins, 1998) described earlier. The WM treatment task, called sequenced exercises for WM, was secondary to a speeded reading aloud treatment task. This task involved sets of grammatical and ungrammatical sentences, organised hierarchically in three levels of difficulty comprising 2, 4 and 6 sentences with 6, 8 and 10 words per sentence respectively. The person read each sentence and judged it for grammatical correctness (processing component). He also had to remember the semantic category of the last word of each sentence and recall it at the end of the set of sentences (storage component). For the two-sentence set, one semantic category had to be recalled, while for the sets of four and six sentences, two and three categories had to be recalled respectively. Presumably, the sentences were different from session to session, although the authors do not provide this information. Mayer and Murray argued that this task resembles several cognitive processes of reading by requiring selective attention to detail and lexical-semantic working memory, which were impaired in their participant.

Most other studies used serial repetition of words, phrases and sentences (Francis et al., 2003; Majerus et al., 2005; Koenig-Bruhin & Studer-Eichenberger, 2007; Kalinyak-Fliszar et al., 2011; Berthier et al., 2014). Francis et al. (2003) used hierarchical repetition tasks that started with two function words, gradually stepped up to two- three- and up to seven-word sentences, before moving on to linguistically more complex sentences (e.g., reversible passive sentences). There were also tasks whereby their participant had to listen to a sentence and match it to a written one, avoiding distracting written sentences. For each treatment session, which was carried out as homework with the assistance of the participant's husband, the sentences were different. A similar treatment protocol in terms of linguistic material and hierarchy was also used by Koenig-Bruhin and Studer-Eichenberger (2007). However, in this study, delayed repetition was also used, varying from 5 to 10-12 seconds between hearing an item and repeating it.

Majerus et al. (2005) and Kalinyak-Fliszar et al. (2011) used words and non-words that had to be repeated immediately and, as the treatment progressed, with a delay which was either filled (Majerus et al., 2005; Kalinyak-Fliszar et al., 2011) or unfilled (Kalinyak-Fliszar et al., 2011). In the delayed recall stage of five seconds, the participant in Majerus et al. had to count backwards from 5 to 1 before repeating the items. The participant in Kalinyak-Fliszar et al. had to repeat the stimuli (words, non-words) at two temporal intervals of 1 and 5 seconds. The 5 second interval was either silent or filled (i.e., the participant and the clinician counted digits in unison). Harris et al. (2014) contrasted two STM treatments: A phonological that involved serial repetition of non-words lists; a semantic (sic) that involved repetition of real words. In the semantic treatment, the participants were also encouraged to think of the meaning of the words.

Berthier et al. (2014) used sentence repetition. These authors contrasted two treatments (with patients who were also receiving drug therapy): Distributed Speech-Language Therapy (DSLT) and Massed Sentence Repetition Therapy (MSRT). DSLT involved standard aphasia treatment tasks such as naming, repetition, sentence completion and conversation among other tasks, carried out by a speech and language therapist. Participants carried out the more intensive STM treatment (MSRT) at home, by listening to pre-recorded sentences and repeating them in the absence of a clinician. Initially, a clinician provided a practice session. There were two sets of 20 sentences (from 2 to 7 words), although it is not clear if the same sentences were used again and again, or different sentences were used. This is important because using the same materials in each session would result in long-term memory retention, rather than STM treatment (Francis et al., 2003; Harris et al., 2014).

We note that repetition is a popular task in STM treatments. We are also aware that repetition is a long-standing treatment method used widely by clinicians. However, the emphasis in STM treatments that use repetition as a treatment method is on the serial recall of words, the words that are used usually vary from session to session, the lists become progressively longer and, in some cases, a delay condition is used, thereby increasing the STM demands of the task.

Murray et al. (2006) used a published programme, the attention processing training programme (APT-II) (Sohlberg, Johnson, Paule, Raskin & Mateer, 2001). This is a structured protocol for treating attentional deficits in people with mild cognitive impairment. Although the programme focuses on attention, many of the tasks involve WM (cf., the link of attention and WM discussed earlier). The tasks aimed to improve hierarchical levels of attention from simple to complex (i.e., sustained, alternating, selective, divided attention). For example, the participant listened to word-lists and then had to identify the items that were round.

Salis (2012) used matching listening span tasks of words. Initially, monosyllabic word lists were used and, as the participant’s ability to identify the order of words in the spoken lists improved, polysyllabic words were used. In each treatment session, different words were used. The rationale behind this approach was to improve spoken sentence comprehension in the person with aphasia. However, because of the fairly severe language comprehension deficit of the participant, a more traditional treatment for sentence comprehension was not deemed appropriate.

The treatment protocols used by Vallat and colleagues (Vallat et al., 2005; Vallat-Azouvi et al., 2014) involved several tasks: Metalinguistic phonological and orthographic processing tasks (reconstitution of words from oral spelling, with or without letter omission, oral spelling, odd or even number of letters in a word, reconstitution of words from syllables, word sorting in alphabetic order, acronyms). There were also visual WM tasks (two- and three-dimensional mental imagery), two n-back tasks (monitoring of words and questions). Finally, there was also a spoken variant of the WM task used by Mayer and Murray (2002) with more levels of difficulty than what Mayer and Murray used.

Critical evaluation: Overall, the studies we reviewed reported improvements on STM/WM tests and, in some cases, other language functions (table 6). In the majority of studies, the researchers sought to address their participants' self-reported difficulties regarding STM/WM functioning, which impacted upon everyday communicative functioning (e.g., Vallat et al., 2005). To date, this is perhaps the most important contribution of STM/WM treatments to person-centred care because the treatment goals addressed the insights and concerns of the person with aphasia as well as their concomitant STM/WM impairments and how these impairments interfered with everyday communication. However, there are several caveats that need to be considered, before adopting the published evidence in clinical practice.

While in all studies the design used was a single-case experimental, in that all studies involved control tasks to isolate treatment effects from other factors and demonstrate treatment effectiveness, the robustness of the designs and consequent threat to internal validity varied. At least two pre-treatment baseline measures (either in language processing or STM/WM measures) were reported by Majerus et al. (2005), Vallat et al. (2005), Murray et al. (2006), Kalinyak-Fliszar et al. (2011), Salis (2012), Vallat-Azouvi et al. (2014). Stable baselines were evident in Vallat et al. (2005), Kalinyak-Fliszar et al. (2011), Salis (2012) and Vallat-Azouvi et al. (2014). In Majerus et al. (2005), the assessment data given for the first baseline phase were limited. In Murray et al. (2006), only one of the two baseline probes showed stability. In the remaining studies, only one pre-treatment baseline in the STM/WM treatment tasks was taken (Mayer & Murray, 2002; Francis et al. 2003; Koenig-Bruhin & Studer-Eichenberger, 2007; Berthier et al., 2014; Harris et al., 2014). Baselines are important because their stability can show if a person improves in response to task exposure as opposed to task treatment, that is, in the presence of treatment feedback.

With a few exceptions (Murray et al., 2006; Harris et al., 2014; Berthier et al., 2014), replications of treatments have not been reported in the literature. In Harris et al. (2014) and Berthier et al. (2014), the same treatments were replicated in at least one other participant within the same study. One of the limitations in Harris et al. (2014) is that of order effects. The person with a phonological STM impairment improved in sentence repetition after phonological STM treatment. The person with a semantic STM impairment improved on sentence anomaly judgement, only after semantic STM treatment and not after phonological STM treatment. While treatment effects were demonstrated in terms of the control tasks (tests of broader cognition), as the authors acknowledge, the semantic STM treatment may have produced the same results if it was given first. Replications are crucial because it is important to know if a treatment can be effective for people with different impairments and different lesions

(Price, Seghier & Leff, 2010). Replications are also important in relation to the clinical setting a reported treatment took place. In Berthier et al. (2014), the STM treatment was self-administered as homework. In Francis et al. (2003) and Harris et al. (2014) parts of the STM treatment were also given as homework. The same is true for part of the treatment by Salis (2012), although this feature was not reported. Homework encourages self-care, motivation, transfer of skill and increases treatment intensity. However, homework can also be detrimental to the internal validity of a study because researchers have very little knowledge as exactly how a person used the tasks. Replications are needed to strengthen the current evidence base, ideally from different researchers to eliminate bias. Publication bias, whereby ineffective treatments tend not to be published in peer-reviewed journals (Peplow, 2014), could also be another reason as to why this literature domain is particularly small.

There are also limitations in reporting maintenance effects at least in one of the measures, either treatment to demonstrate treatment effects, or outcome measures to demonstrate maintenance of treatment effects. Maintenance measurements were evident in Mayer and Murray (2002), Murray et el. (2006), Kalinyak-Fliszar et al. (2011) and Vallat-Azouvi et al. (2014).

Two studies combined STM/WM treatments with more mainstream treatment approaches (Mayer & Murray, 2002; Berthier et al., 2014). It could be argued that the STM/WM treatments may have been more beneficial in improving the participants' communicative functioning. However, because the memory treatments were given alongside other treatments (in Mayer and Murray a reading aloud treatment; in Berthier et al. a mixture of mainstream treatment tasks), it is unclear if the STM/WM treatment components were more effective than the mainstream treatments. Similarly, several treatment tasks were used by Vallat and colleagues (Vallat et al., 2005; Vallat-Azouvi et al., 2014). While overall improvements were reported, the comparative effectiveness of the tasks is unclear. More evidence is needed in order to contrast the relative benefits of STM/WM treatments in relation to more mainstream language treatments.

Only two studies reported treatment fidelity (Mayer & Murray, 2002; Kalinyak-Fliszar et al., 2011). Treatment fidelity ensures that the treatment protocol was applied consistently across sessions. One study (Salis, 2012) reported ‘blinded’ assessment at the end of treatment for the two main outcome measures. However, treatment fidelity was not addressed. Ideally, pre- and post-treatment assessment should be ‘blinded’ and be carried out by a person not involved and not aware of the focus of the treatment.

In terms of STM/WM treatment outcome measures, all studies used at least one STM or WM measure to evaluate treatment effects. Based on the task descriptions provided, these measures (e.g., forward and backward digit span, listening span) have construct validity in that they measure STM and WM abilities (the constructs of interest). Only a few studies used at least one STM/WM measure with known test-retest reliability. These were either the digit span of earlier versions of the Wechsler batteries (Mayer & Murray, 2002; Francis et al., 2003; Murray et al., 2006; Salis, 2012), or the listening span task (Murray et al., 2006). In the remaining of studies, the STM/WM outcome measures were tasks with unknown test-retest reliability.

Finally, measures of everyday functioning, documenting the clinical significance of the treatments were only reported by Vallat et al. (2005), Murray et al. (2006) and Vallat-Azouvi et al. (2014). In Vallat et al. (2005) their participant improved in both measures of everyday functioning (verbal communication and WM questionnaires).

In Vallat-Azouvi et al. (2014) performance in WM functioning improved as evidenced by a WM questionnaire. Using the CETI (Lomas, Pickard, Bester, Elbard & Zoghaib, 1989), Murray and colleagues did not find a difference before and after treatment. The CETI was developed to capture perceptions of communicative functioning in aphasia and may not have been the most sensitive measure to capture the treatment-induced changes in Murray et al. (2006).

Conclusions

In this paper, we argued for the adoption of STM and WM treatments in current aphasia management for two main reasons: First, the foundational link between STM and WM in language; second, the prevalence of STM and WM impairments in stroke aphasia. We described several tests to assess STM and WM impairments, with a view to broadening current assessment practices. We also provided a critical review of published treatment studies. Based on methodological concerns, it is unclear as to why some treatments have been shown to improve STM or WM and, importantly, to transfer/generalisation effects in other language functions. To date, it can be concluded that there is no evidence of absence of treatment effects attributed to STM/WM treatments, but an absence of robust evidence. Previously, we discussed the close links between STM/WM performance and language processing. These links lead to the hypothesis that treatment of STM/WM would improve language functions and also STM/WM. According to Howard and Hatfield (1987), a key principle in treatment planning is not to establish beyond all doubt the underlying impairment. All that is required is a hypothesis sufficiently detailed to motivate therapy. The theoretical rationale for incorporating STM/WM treatment protocols in clinical practice is considerably stronger than the quality of treatment effectiveness. At the very least, clinicians should assess STM and WM memory functioning in stroke aphasia because, as Martin and Reilly (2012) put it, the content of STM (and WM) is language. Practice-based evidence, like evidence-based practice calls for high quality, patient-centred care. Having discussed theory, assessment and treatment descriptions, together with a critical evaluation of treatments, we would like to encourage clinicians to strengthen the evidence of this novel and, likely, promising avenue of clinical research.

References

BADDELEY, A. 2012, Working memory: Theories, models, and controversies. Annual Review of Psychology, 63, 1-29.

BEESON, P. M., BAYLES, K.A., RUBENS, A.B., and KASZNIAK, A. W. 1993, Memory impairment and executive control in individuals with stroke-induced aphasia. Brain and Language, 45, 253-275.

BERTHIER, M., L., DAVILA, G., GREEN-HEREDIA, C., MORENO TORRES, I., DE MIER, R., and DE-TORRES, I., 2014, Massed sentence repetition training can augment and speed up recovery of speech production deficits in patients with chronic conduction aphasia receiving donepezil treatment. Aphasiology, 28, 188-218.

BRADY, M., KELLY, H., GODWIN, J., and ENDERBY, P., 2012, Speech and language therapy for aphasia following stroke. Cochrane Database of Systematic Reviews, 16; 5, CD000425.

CASPARI, I., PARKINSON, S. R., LAPOINTE, L. L. and KATZ, R. C. 1998, Working memory and aphasia. Brain and Cognition, 37, 205-23.

CHRISTENSEN, S. C., and WRIGHT, H. H., 2010, Verbal and non-verbal working memory in aphasia: What three n-back tasks reveal. Aphasiology, 24, 752-762.

COWAN, N., 2008, What are the differences between long-term, short-term, and working memory? Progress in Brain Research, 169, 323-338.

CROCKET, D., CLARK, C., SPREEN, S., and KLONOFF, H., 1981, Severity of impairment or specific types of aphasia: An empirical investigation. Cortex, 17, 83-96.

DEDE, G., RICCA, M., KNILANS, J., and TRUBL, B., 2014, Construct validity and reliability of working memory tasks for people with aphasia. Aphasiology, 28, 692-712.

DELL, G. S., SCHWARTZ, M. F., MARTIN, N., SAFFRAN, E. M., and CAGNON, D., 1997, Lexical access in normal and aphasic speakers. Psychological Review, 104,801-838.

DE RENZI, E., and VIGNOLO, L. A., 1962, The Token Test: A sensitive test to detect receptive disturbances in aphasics. Brain, 85, 665-678.

DOBINSON, C., and WREN, Y. (2013). Creating Practice-based Evidence: A guide for SLTs (Guildford, UK: J & R Press).

GOODGLASS, H., BERKO GLEASON, J., and HYDE, M. R., 1970, Some dimensions of auditory language comprehension in aphasia. Journal of Speech and Hearing Research, 13, 595-606.

FRANCIS, D. R., CLARK, N., and HUMPHREYS, G. W., 2003, The treatment of an auditory working memory deficit and the implications for sentence comprehension abilities in mild "receptive" aphasia. Aphasiology, 17, 723-750.

FINGER, S., 1994, Origins of Neurosciences: A history of explorations into brain function (Oxford: Oxford University Press).

GOODGLASS, H., KAPLAN, E., and BARRESI, B., 2001, The Assessment of Aphasia and Related Disorders (3rd ed.) (Baltimore: Lippincott Williams & Wilkins).

GORDON, W. P., 1983, Memory disorders in aphasia I: Auditory immediate recall. Neuropsychologia, 21, 325-339.

GREENLAGH, T. 2014, How to Read a Paper: The basics of evidence-based medicine (Chichester, UK: John Wiley).

HARRIS, L. H., OLSON, A., and HUMPHREYS, G., 2014, The link between STM and sentence comprehension: A neuropsychological rehabilitation study. Neuropsychological Rehabilitation: An International Journal, 24, 678-720.

HASHER, L., LUSTIG, C., & ZACKS, R. T., 2007, Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, A., & J. Towse (Eds.), Variation in working memory (pp. 227-249) (New York: Oxford University Press).

HOWARD, D., and HATFIELD, F. M., 1987, Aphasia Therapy: Historical and contemporary issues (London: Lawrence Erlbaum Associates).

HOWARD, D., and NICKELS, L., 2005, Separating input and output phonology: Semantic, phonological, and orthographic effects in short-term memory impairment. Cognitive Neuropsychology, 22, 42-77.

INTERCOLLEGIATE STROKE WORKING PARTY, 2012, National Clinical Guideline for Stroke (4th ed) (London: Royal College of Physicians).

JAMES, W., 1890, The Principles of Psychology (New York: Henry Holt).

KALINYAK-FLISZAR, M. KOHEN, F., and MARTIN, N., 2011, Remediation of language processing in aphasia: Improving activation and maintenance of linguistic representations in (verbal) short-term memory. Aphasiology, 25, 1095-1131.

KERTESZ, A., 1982, Western Aphasia Battery (New York: Grune & Stratton).

KERTESZ, A., 2006, Western Aphasia Battery - Revised (London: Pearson).

KOENIG-BRUHIN, M., and STUDER-EICHENBERGER, F., 2007, Therapy of short-term memory disorders in fluent aphasia: a single case study. Aphasiology, 21, 448-458.

LAPOINTE, L. L., and HORNER, J., 1978, Reading Comprehension Battery for Aphasia (Austin, TX: Pro-Ed).

LAURES-GORE, J., SHISHLER-MARSHAL, R., and VERNER, E., 2011, Performance of individuals with left-hemisphere stroke and aphasia and individuals with right brain damage on forward and backward digit span tasks. Aphasiology, 25, 43-56.

LEFF, A. P., SCHOFIELD, T. M., CRINION, J. T., SEGHIER, M. L., GROGAN, A., GREEN, D. W., and PRICE, C. J., 2009, The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: Evidence from 210 patients with stroke. Brain, 132, 3401-3410.

LEHMAN, M. T., and TOMPKINS, C. A., 1998, Reliability and validity of an auditory working memory measure: Data from elderly and right-hemisphere damaged adults. Aphasiology, 12, 771-785.

LESSER, R., 1976, Verbal and non-verbal memory components in the Token Test. Neuropsychologia, 14, 79-85.

LOMAS, J., PICKARD, L., BESTER, S., ELBARD, H., FINLAYSON, A., and ZOGHAIB, C., 1989, The communicative effectiveness index: Development and psychometric evaluation of a functional communication measure for adult aphasia. Journal of Speech and Hearing Disorders, 54, 113-124.

MAJERUS, S., VAN DER KAA, M-A., RENARD, C., VAN DER LINDEN, M., and PONCELET, M., 2005, Treating verbal short-term memory deficits by increasing the duration of temporary phonological representations: A case study. Brain & Language, 95, 174-175.

MARTIN, N., and AYALA, J. 2004, Measurements of auditory-verbal STM span in aphasia: effects of item, task, and lexical impairment. Brain and Language, 89, 464-

83.

MARTIN, N., and GUPTA, P., 2004, Exploring the relationship between word processing and verbal short-term memory: Evidence from association and dissociations. Cognitive Neuropsychology, 21, 213-228.

MARTIN, N., and SAFFRAN, E. N., 1999, Effects of word processing and short-term memory deficits on verbal learning: Evidence from aphasia. International Journal of Psychology, 34, 330- 346.

MARTIN, N., SAFFRAN, E. M., and DELL, G. S. 1996, Recovery in deep dysphasia: Evidence for a relation between auditory-verbal STM capacity and lexical errors in repetition. Brain and Language, 52, 83-113.

MARTIN, N., and REILLY, J., 2012, Short-term/working memory impairments in aphasia: Data, models, and their application to aphasia rehabilitation. Aphasiology, 26, 253-257.

MARTIN, R. C., and ALLEN, C. M., 2008, A disorder of executive function and its role in language processing. Seminars in Speech and Language, 29, 201-210.

MAYER, J. F., and MURRAY, L. L., 2002, Approaches to the treatment of alexia in chronic aphasia. Aphasiology, 16, 727-743.

MAYER, J. F., and MURRAY, L. L., 2012, Measuring working memory deficits in aphasia. Journal of Communication Disorders, 45, 325-339.

McKEVITT, C., FUDGE, N., REDFERN, J., SHELDENKAR, A., CRICHTON, S., RUDD, A. R., FORSTER, F., YOUNG, J., NAZARETH, I., SILVER, L. E., ROTHWELL, M., and WOLFE, C. D. A., 2011, Self-reported long-term needs after stroke. Stroke, 42, 1398-1403.

McNEIL, M. R., and PRESCOTT, T. E., 1978, The Revised Token Test (Baltimore, MD: University Park Press).

MURRAY, L. L., 2012, Direct and indirect treatment approaches for addressing short-term or working memory deficits in aphasia. Aphasiology, 26, 317-337.

MURRAY, L. L., KEETON, R. J., and KARCHER, L., 2006, Treating attention in mild aphasia: Evaluation of attention process training-II. Journal of Communication Disorders, 39, 37-61.

PEPLOW, M., 2014, Social sciences suffer from severe publication bias: Survey finds that ‘null results’ rarely see the light of the day. Nature, doi:10.1038/nature.2014.15787

POTAGAS, C., KASSELIMIS, D., and EVDOKIMIDIS, I., 2011, Short-term and working memory impairments in aphasia. Neuropsychologia, 49, 2874-2878.

PRICE, C. J., SEGHIER, M. L., and LEFF, A. P., 2010, Predicting language outcome and recovery after stroke: The PLORAS system. Nature Reviews Neurology, 6, 202-210.

RANDOLPH, C., 2012, Repeatable Battery for the Assessment of Neuropsychological Status (London, UK: Pearson).

REYNOLDS, C. R., and VORESS, J. K., 2007, The Test of Memory and Learning (TOMAL-2) (Austin, TX: Pro-Ed).

ROGERS, M. A., and STORKEL, H. L., 1999, Planning speech one syllable at a time: The reduced buffer capacity hypothesis of apraxia of speech. Aphasiology, 13, 793-805.

ROYAL COLLEGE OF SPEECH AND LANGUAGE THERAPISTS, 2005, Clinical guidelines (Bicester, UK: Speechmark).

SALIS, C., 2012, Short-term memory treatment: Patterns of learning and generalisation to sentence comprehension in a person with aphasia. Neuropsychological Rehabilitation: An International Journal, 22, 428-448.

SCHUELL, H., JENKINS, J. J., and JIMENEZ-PABON, E., 1964, Aphasia in Adults: Diagnosis, prognosis and treatment (London: Harper Row).

SNAPHAAN, L., and DE LEEUW, F. E., 2007, Post stroke memory function in non demented patients: A systematic review on frequency and neuroimaging correlates. Stroke, 38,198-203.

SOHLBERG, M., JOHNSON, L., PAULE, L., RASKIN, S., and MATEER, C., 2001, Attention Process Training: A Program to Address Attentional Deficits for Persons with Mild Cognitive Dysfunction (an upper extension of the APT-I) (2nd ed) (Wake Forest, NC: Lash & Associates Publishing/Training).

STUART G. P., and HULME, C., 2009, Lexical and semantic influences on immediate serial recall: A role for redintegration. In A. Thorn and M. Page (eds), Interactions between Short-term and Long-term Memory in the Verbal Domain (Hove: UK: Psychology Press), pp. 157-176.

SWINBURN, K., PORTER, G., & HOWARD, D., 2005, The Comprehensive Aphasia Test (Hove, UK: Psychology Press).

TESAK, J., and CODE, C., 2008, Milestones in the History of Aphasia: Theories and Protagonists (Hove, UK: Psychology Press).

TOMPKINS, C. A., BLOISE, C. G. R., TIMKO, M. L., and BAUMGAENTNER, A., 1994, Working memory and inference revision in brain damaged and normally aging adults. Journal of Speech and Hearing Research, 37, 896-912.

VALLAT, C. A. P., HARDISSON, H., MEFFERT, R., TESSIER, C., and PRADAT-DIEHL, P., 2005, Rehabilitation of verbal working memory after left hemisphere stroke. Brain Injury, 19, 1157-64.

VALLAT-AZOUVI, C., PRADAT-DIEHL, P., and AZOUVI, P., 2014, Modularity in rehabilitation of working memory: A single-case study. Neuropsychological Rehabilitation: An International Journal, 24, 220-237.

Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Jennifer L. Beaumont, J. L., Kisala, P. A., Mungas, D., Conway, K., and Gershon R., 2014, NIH toolbox cognition battery (NIHTB-CB): List sorting test to measure working memory. Journal of the International Neuropsychological Society, 20, 599-610.

WARRINGTON, E. K., and SHALLICE, T., 1969, The selective impairment of auditory-verbal short-term memory. Brain, 92, 885–896.

WATERS, G. S., and CAPLAN, D., 2003, The reliability and stability of verbal working memory measures. Behavior Research Methods, Instruments and Computers, 35, 550-564.

WHITWORTH, A., WEBSTER, J., and HOWARD, D., 2013, A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia: A clinician's guide (London: Psychology Press).

WECHSLER, D., 2010, Wechsler Adult Intelligence Scale IV-UK (London, UK: Pearson).

WECHSLER, D., 2010, Wechsler Memory Scale IV-UK (London, UK: Pearson).