CORTEX XXX (2018) 1–14

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Research report

How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM)

Jeremy J. Tree[*] *and David Playfoot*

Department of Psychology, Swansea University, Singleton Park, Swansea, UK

ARTICLE INFO

Article history:
Received 1 December 2017
Reviewed 12 February 2018
Revised 1 April 2018
Accepted 20 August 2018
Published online xxx

Keywords:
Conduction aphasia
Short-term memory
Immediate serial recall

ABSTRACT

In the field of cognitive neuropsychology of phonological short-term memory (pSTM), a key debate surrounds the issue of how impairment on tasks deemed to tap this system imply a dissociable phonological input and output buffer system, with the implication that impairments can be fractionated across disruption to separate functional components (Nickels, Howard & Best, 1997). This study presents CT, a conduction aphasic who showed no impairment on basic auditory discrimination tasks, but had very poor nonword repetition. Clear-cut examples of such cases are very rare (see Jacquemot, Dupoux & Bachoud-Levi, 2007), and we interpret the case with reference to a pSTM model that includes input and output buffers. The dissociation between performance on auditory phonological tasks and visual phonological tasks we interpret as consistent with disruption to the link from input buffer to output buffer without concurrent damage to connections from output to input. Previous research has also shown that patients with impairments of pSTM can make visual confusions with orthographically presented items in tasks seeking to tap this mechanism (Warrington & Shallice, 1972), which might stem from having an incomplete pSTM loop. In light of this we examined whether CT's ability on tests of ISR was affected by visual orthographic similarity among list items, and this is indeed what we observed. On balance then, CT's overall profile is considered best interpreted with respect to a dual buffer pSTM model (e.g., Vallar & Papagno, 2002).

Patients with a condition known as conduction aphasia often present with a severe impairment in speech repetition that shows a lexical advantage; with word repetition better than nonword repetition (Shallice & Warrington, 1977). Such cases often have fluent speech production and preserved auditory comprehension (Bartha & Benke, 2003; Goodglass, 1992). Conduction aphasia is therefore a particularly interesting disorder for theories concerning the nature of the cognitive systems that deal with phonological input (speech perception) and output (speech production) codes (and the storage/manipulation thereof), which we will call the phonological short-term memory system (pSTM). Research on this system with respect to both normal and neuropsychological populations has a long and varied history, and theoretical models

---

[*] *Corresponding author*. Department of Psychology, Swansea University, Singleton Park, Swansea, SA2 8PP, UK.
E-mail addresses: j.tree@swansea.ac.uk (J.J. Tree), d.playfoot@shu.ac.uk (D. Playfoot).

https://doi.org/10.1016/j.cortex.2018.08.008

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

of the functional architecture of this system have debated how things might be organised – perhaps one 'classical' model of this system is that suggested by Monsell (1987 outlined below). This model favours the inclusion of 'buffer' systems and dissociable functional components for which specific damage results in predictable selective patterns of impairment in neuropsychological cases such as those with conduction aphasia. To foreshadow our work presented here, we will describe a conduction aphasic patient, CT, and frame his performance with respect to this dual buffer model of phonological short-term memory (pSTM) proposed by Monsell (1987). We hope to show that not only is a classical 'buffer' model of pSTM still relevant to work such as ours, but it constitutes an important tool for understanding the pattern of impairment that CT presents with, and provided us with the impetus to undertake specific work relating to assessment of other forms of non-auditory based processing (i.e., orthographic processing and memory tasks).

1. A dual-buffer model of pSTM – implications for interpreting neuropsychological cases

The motivation for dual buffer models has classically been provided by neuropsychological case studies to make sense of two contrasting behavioural presentations. On the one hand, there are patients who have presented with impairments in phonological input processing without any accompanying deficits in spontaneous output processing (e.g., Allport, 1984; Romani, 1992). On the other hand, there are patients who are reported with problems in processing phonological output, but not input (e.g., Martin, Shelton & Yaffee, 1994). The implication being that these dissociative patterns of impairment, are not easily reconciled with models of pSTM that propose only a single buffer; since such a model would assume a more generalised pattern of impairment across both phonological input and output tasks. Somewhat motivated by the interpretation of this neuropsychological evidence by many in the field (e.g., Howard & Franklin, 1988), that it indicates that it is possible to lose some aspects of pSTM ability without being universally impaired, Monsell (1987) proposed a model of pSTM containing separate input and output buffers (Fig. 1). In this model (see also Tree & Kay, 2015; Vallar & Papagno, 2002), auditory input enters the input buffer directly, which allows for short-term storage (e.g., a telephone number you hear before writing it down) and subsequent access to semantic memory (for speech comprehension). The model proposes a phonological output buffer, which stores phoneme level information and is the point of entry to pSTM for visual input (as you convert orthographic input to phonological output). It is worth noting that the previous work has divided conduction aphasia into *reproduction* and *repetition* subtypes (e.g., Shallice & Warrington, 1977, see also; Gvion & Friedmann, 2012). Both of these disorders have impaired repetition performance as the cardinal feature, but reproduction conduction aphasia is also characterised by the presence of paraphasias and neologisms which are absent in the repetition subtype. The explanations for the performance of these cases by previous authors have made specific reference to differential damage to the input buffer (for repetition conduction aphasia, Bartha & Benke, 2003; Butterworth, 1992; Howard & Nickels, 2005; Martin & Breedin, 1992; Martin, Shelton, et al., 1994; Shallice, Rumiati & Zadini, 2000) or output buffer (for reproduction conduction aphasia, Code, Tree & Dawe, 2009; Franklin, Buerk, & Howard, 2002; Kohn, 1992; Kohn & Smith, 1994). In addition, Monsell's (1987) model contains a rehearsal mechanism which consists of a link from input buffer to output buffer and a separate link from output to input, the combination of which acts to prevent trace decay (equivalent to when you might repeat that phone number to yourself to keep it in mind). We would argue that in addition to the suggestion of multiple buffers, this inclusion of multiple pathways is also a key feature of the model because it is theoretically possible for one of these links to become damaged while the other is intact. To again foreshadow our subsequent discussion of our neuropsychological case (CT), we will also argue that 'dual buffer' models such as this, can readily account for a key feature of his pattern of impairment, namely very poor nonword repetition despite normal speech comprehension and articulation. Because cases such as his are able to understand speech (i.e., their phonological *input* codes are intact) and have relatively few problems in spontaneous speech production (i.e., their phonological *output* codes are intact), we would interpret the overall pattern of impairment as perhaps best captured by a disruption of the connection transferring phonological representation between the input and output codes in pSTM.

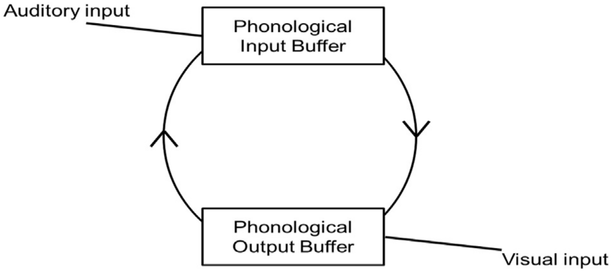


Fig. 1 – A sketch of Monsell's (1987) model of phonological short term memory.

In a now classic study, Nickels, Howard & Best (1997), with reference to a dual-buffer pSTM model, provided a key narrative with which to interpret patterns of impairment in patients as reflecting disruption to the link from input to output, or the link from output to input. Nickels et al. (1997) tested aphasic patients on three key verbal short-term memory tasks: (a) auditory rhyme judgement (e.g., CAT-MAT do they rhyme?), (b) visual (written word) rhyme judgement, and (c) another written word phonological task, homophone judgement (e.g., SON-SUN, do they sound the same?), The authors argued that these three tasks differed in two important ways: (1) whether they required the transference of orthographic to phonological codes (and thus direct access to the output buffer), which is true for (b) and (c) but not (a), and (2) whether they required the segmentation, storage and comparison of phonological codes, which is true of (a) and (b) but not (c). Importantly, in a dual-buffer pSTM model such as Monsell's (1987), there is a clear prediction that is made with respect to these different tasks, namely that impaired

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

performance across tasks can dissociate (i.e., patients can be selectively impaired at one with the others relatively spared). For example, Monsell's (1987) model allows for the possibility that homophone judgement and auditory rhyme judgement are preserved while written rhyme judgement is impaired – this would occur when the connection from the output buffer to the input buffer is disrupted. To understand this, note that spoken stimuli enters pSTM at the input buffer, which allows for the segmentation/comparator processes (see 2 above) that enable auditory rhyme judgements to remain unhindered. Meanwhile, orthographic-to-phonological information enters pSTM at the output buffer. Nickels et al. (1997) argued that for tasks such as homophone judgement (that do not require segmentation) storage in the output buffer is sufficient, and thus this ability is preserved. However, successful written word rhyme judgement requires access to the input buffer (to enable the additional processing also required for auditory rhyme judgement) and thus if the output-input link is damaged, impaired performance will emerge. In this example, not only do we see a means by which *selective* impairment of a particular phonological task can occur, but the same model can also explain other specific patterns of impairment.

Monsell's (1987) model also suggests that damage to the input-output link will have a detrimental effect on pSTM span tasks, (since rehearsal is not possible) but will be of no consequence to homophone or rhyme judgement (both auditory or visual) performance. Importantly, none of the cases described in Nickels et al. (1997) showed the general impairment predicted by single buffer models, but the performance of every patient could be accommodated within Monsell's (1987) framework. For the present work, we use Monsell's (1987) theoretical framework to probe the nature of the impairment in our case CT. We used a number of different tasks (such as visual and auditory rhyme judgement, and homophone judgement as described above) for which a specific pattern of task impairment could be interpreted as being attributed to a key functional component. In so doing, we hope to demonstrate not only the *validity* of such a dual-buffer pSTM model, but also the *utility* of such a model in helping understand the nature of the functional impairment that can account for cases such as CT (both of which we would argue are classic objectives of a great many studies published in the journal *Cortex*).

In Fig. 2, below, we map each of our key neuropsychological tasks to components within a dual-buffer pSTM model – this is in line with the work of Nickels et al. (1997) and aims to show how we might interpret specific task impairment in each and every case. We argue that the purpose of administering such a variety of tests with reference to a well-specified functional model such as this allows for the interpretation of any emerging dissociative patterns of impairment *a priori*. For example, let us consider a cardinal feature of conduction aphasia – that word repetition performance is usually better than non-word repetition performance (Shallice & Warrington, 1977). This has been explained by suggesting that word (unlike nonword) repetition benefits from the access to stored phonological information about familiar phonological forms (i.e., representations stored in the input and output lexicons presented in Fig. 2). Thus on reflection of the model in Fig. 2, a word can be "repeated" by either, a) mapping from auditory input phonology directly to spoken output phonology in short term memory (the input-output buffer link) or b) recognising the word that has been presented and retrieving its phonology indirectly from long term memory (the so called 'lexical' repetition route – see McCarthy & Warrington, 1984, Hickok & Poeppel, 2004). On the other hand, non-word repetition is only achievable via the direct input-output buffer pathway. If we assume a patient has an impairment of the direct pathway (arrow a in Fig. 2), this case should present with a severe disruption of nonword repetition, with relatively preserved word repetition (so often reported in conduction aphasic cases such as CT discussed here). However, under the same model, if either the input or output lexicon is also damaged, then the lexicality advantage would likely be absent.

A particularly relevant case in the literature is that of FA (Jacquemot, Dupoux & Bachoud-Levi, 2007). FA was a conduction aphasic who showed no deficits in speech perception at either the phonological or the semantic level (as determined by a minimal pairs discrimination task and spoken word to picture matching, respectively). She *was* impaired in speech production measured by picture naming, word and nonword repetition and reading aloud. The authors argued that the errors in word repetition and reading FA made were due to a slight global word production deficit; that the patient's nonword repetition performance was considerably worse and consisted of qualitatively different types of errors was interpreted as evidence of a deficient direct repetition route. Following logic similar to Nickels et al. (1997), Jacquemot et al. (2007) went on to describe FA's performance on written rhyme judgement and a second task in which the patient was presented with a picture and spoken item (either word or nonword) and asked whether the name of the picture and the sound they had heard were a rhyming pair. They determined that FA was impaired relative to controls on both of these phonological tasks. The authors argued that this could be accommodated in a model which proposed a) separate phonological input and output codes and b) separate conversion links from phonological input to output, and in the opposite direction. They suggested that FA had an impairment to the input-output conversion, but no impairment for the output-input conversion.

2. Dual buffer models of pSTM – implications for patterns of immediate serial recall impairment

Under the assumption that the performance of our case matched the predictions based on Monsell's (1987) model and the findings of Nickels et al. (1997), we pre-emptively considered the impact of disruption to just this half of the rehearsal loop ('a' in Fig. 2) on immediate serial recall (ISR) performance in auditory and visual modalities. In ISR tasks, participants are exposed to lists of stimuli and are instructed to remember the items in the order they were presented. Responses are scored as correct only if they were recalled in the appropriate position within the list. As a general rule, auditory immediate serial recall (AISR) tends to be generally better than visual immediate serial recall (VISR) in healthy participants. In

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

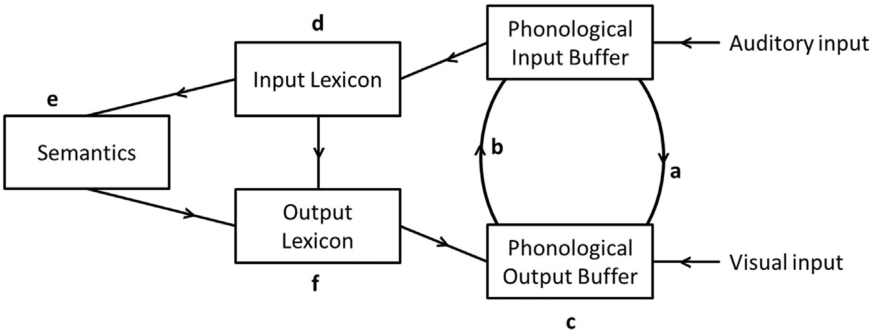


Fig. 2 – The integration of auditory verbal short term memory with broader language processes. Letters represent the processes particularly assessed by the tests administered to CT. a = input-output link assessed by non-word repetition, b = output-input link assessed by written rhyme judgement, c = phonological output buffer assessed by homophone judgement, d = lexical memory assessed by auditory lexical decision, e = semantic processing assessed by Pyramids and Palm Trees (Howard & Patterson, 1992), f = generation of phonological output from non-verbal input assessed by Graded Naming Test (McKenna & Warrington, 1983).

particular, the auditory modality shows the biggest advantage at the recency (i.e., final items) part of the serial position curve (Conrad & Hull, 1968). However, in patients with pSTM deficits, this is very rarely the case. In fact, visual recall tends to be rather better than auditory recall (Best & Howard, 2005; Howard, 1995; patient HB, Howard & Nickels, 2005; Shallice & Vallar, 1990; Warrington & Shallice, 1969; 1972). Warrington and Shallice (1972), for example, reported several experiments that demonstrated that the short-term memory performance of their patient, KF, was worse in the auditory modality than in written presentation, and that forgetting of auditory information was also more rapid. This general pattern of better performance in visual over auditory presentation is curious. If you assume that success at ISR requires phonological information to be stored in a 'buffer' and rehearsed via a 'loop', and that the same 'code' tends to underpin both types of tasks, you might also assume equivalence in performance impairment regardless of modality of presentation. Further, it is possible to assume that given the additional computational demands of a 'translation' process (orthographic-to-phonological recording), written material should always be *worse* than auditory material. If we take the perspective of Monsell's (1987) dual buffer model, then this pattern could be easily explained. In typically functioning participants, input to short-term memory is rehearsed via intact input-output and output-input links. Auditory information is recalled more efficiently than visual information because of the additional processing cost that is incurred to convert the orthography to phonological codes. In patients with an impairment of the link between input-output buffers, such disruption will disproportionately affect auditory versus written stimuli (as seen in cases such as Best & Howard, 2005).

Under this assumption, some of the classic effects observed in short term memory tasks become intriguing in the context of neuropsychological impairment. One such phenomenon is the phonological similarity effect (e.g., Conrad & Hull, 1964; Copeland & Radvansky, 2001). This refers to the finding that memory is less accurate when all of the items in a stimulus list sound similar than when items are phonologically distinct. For example, participants make a greater number of errors when recalling from a list that includes words like *bat*, *mat* and *cat*, than they do when recalling from a list that includes *dirt*, *cup* and *book*. This finding has traditionally been attributed to a failure to discriminate between the articulatory codes that are used to store the words in the phonological input buffer (e.g., Vallar & Baddeley, 1984). Interestingly, the phonological similarity effect has been observed irrespective of whether the lists are presented aurally or in written form when testing healthy participants (e.g., Besner & Davelaar, 1982). Of course, within the framework of the models of pSTM we have already described, this is not surprising because the orthographic code is translated into phonological code in order to be rehearsed. Once the written form has been converted into phonemes, the phonological similarity effect can occur in the same way as for auditory presentation. Again, in neuropsychological cases of short-term memory impairment, however, the picture regarding the phonological similarity effect can be somewhat different. Vallar, Di Betta and Silveri (1997) suggest that, when an individual has a deficit in phonological memory span, they are unlikely to use sub-vocal rehearsal because they do not have the necessary resources available to them. If this is the case, then distinct patterns of performance should be observed in auditory versus visual serial recall tasks. Vallar et al. (1997) argued that the phonological similarity effect could be apparent in cases with pSTM deficits in spoken, but *not* written, presentation. As we described earlier, in the models of pSTM offered by both Baddeley (1986) and Monsell (1987), spoken input enters the phonological input buffer whereas visual input reaches the phonological output buffer first. The phonological similarity of spoken stimuli is therefore immediately detrimental to memory performance in the input buffer. In the written modality, though, there needs to be a recoding stage and the resulting phonemic information must be passed around the rehearsal loop before any confusion can be caused by the phonological similarity among the items. If either the recoding process or the passage of the phonological code to the input buffer is impeded then the phonological similarity effect should be abolished for written material. Indeed, one of the cases described in Vallar et al.'s (1997) paper, TO, showed precisely this pattern of performance – a phonological similarity effect in auditory serial recall but not in visual serial recall. To foreshadow our work again, given neuropsychological patients with damage to pSTM have been shown to perform differently in visual versus auditory presentation of the same task, we explored this issue in our assessment of our case CT.

3. Impairments of pSTM – implications for stored visual (orthographic) representations

The final part of our exploration of CT's pSTM performance stems from another intriguing case, reported by Best and Howard (2005) which speaks to the issue of modality-specific performance in pSTM tasks. Best and Howard (2005) re-examined the performance of their phonological dyslexic case MJK, who consistently showed superior performance on visual versus auditory digit span tasks. They examined the phonological similarity effect across modalities and found the effect was present in the aural but not visual domain. They reasoned that MJK was likely using a phonological coding system similar to non-impaired participants in recalling spoken input. The same system was not being used when the presentation was visual (much as was suggested by Vallar et al., 1997). That is, MJK was argued not to be recoding written items to phonological codes for recall. However, given that MJK's recall for visual presentation was better overall than for auditory stimuli, Best and Howard (2005) argued that she must be using visual coding for these tasks instead. Indeed, Best and Howard reported that visual confusion errors occurred in MJK's recall performance providing tentative evidence for a *visually based* encoding strategy for VISR which boosts performance over AISR and makes performance vulnerable to visual confusability (in effect a corollary of the phonological similarity effect). Again, the work of Warrington and Shallice (1972) is of interest here. KF made a number of "visual" errors in immediate serial recall; a pattern that was not seen for auditory presented items. Warrington and Shallice (1972) argued that this was evidence for the use of a visual code in short-term memory.

It appears, therefore, that phonological and visual processing may interact in short term memory tasks (see also Tree, Longmore, Majerus & Evans, 2011). Tree, Longmore and Besner (2011) demonstrated that visual orthographic processing may be emphasised when pSTM is disrupted using *articulatory suppression* in undergraduate participants – this refers to the requirement that participants repeat a single word or phoneme (e.g., *the*) over and over again while they are learning the list of stimuli – argued to disrupt the phonological rehearsal process (see Levy, 1971). In other words, when healthy participants do a pSTM task under articulatory suppression there is an emphasis on the "visual information" suggested by Baddeley (1986), because of disruption to the rehearsal of the typically utilised phonological code. The important point is that there may be a convergence of performance in this case, that is in the context of disruption to the phonological rehearsal system (whether as a result of brain injury or articulatory suppression) an emerging shift in emphasis occurs for visually presented information; a shift toward using visually based information for retrieval. As a consequence, under articulatory suppression, visually presented items are precluded from entering the loop and hence the phonological similarity effect disappears – there is no longer a significant detriment for words that sound the same versus those that are phonologically distinct. Articulatory suppression does not, however, eliminate the phonological similarity effect for aurally presented lists (see Levy, 1971).

Thus in the context of the present work, we suggest that it is possible that visual coding strategies may also be emphasised in conduction aphasia cases, like CT, for precisely the same reason. The incomplete (or ineffective) rehearsal loop necessitates that different input modalities require separate coding strategies within verbal short term memory, a visual code for written stimuli; a phonological code for spoken. Hence it is plausible that visual similarity between to-be-recalled items in a serial recall task will be of greater detriment to CT than it is to participants with a functional pSTM – and this issue was also explored in the current study.

4. Method

4.1. *Case description*

CT is a 60 year old right handed male. He was educated to university level and had been a lawyer up until he suffered a stroke 5 years before the start of our testing. A CT scan administered a year before our testing determined a lesion to his left parietal lobe (see Fig. 3).

Testing by a NHS audiology clinic had determined that his hearing was entirely normal (hence any deficits in phonological processing were not a consequence of hearing impairment), and his vision was corrected to normal (again indicating that deficits observed were not because of a sensory problem). His speech was well-formed, and at the normal rate without evidence of spontaneous speech errors. Initial interview determined there was no evidence of comprehension problems and day-to-day memory was reported as being largely normal. A speech and language therapist report written in the period between CT's stroke and our first meeting indicated that his single word repetition was poor, and the report concluded his profile was consistent with a conduction aphasia presentation, although details were not that elaborated. By the time of our first session with CT it was clear that his word repetition performance had recovered substantially (though it was still not fully intact - see below), and despite normal word and non-word reading, he showed a marked impairment in non-word repetition. This striking latter impairment prompted the further investigations reported in this paper.

4.2. *Materials*

We administered a variety of standardised tests that were specifically selected to assess CT's general cognitive function, comprehension or semantic processing, reading, repetition ability, and phonological processing. A large number of the tests we administered were drawn from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay,

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

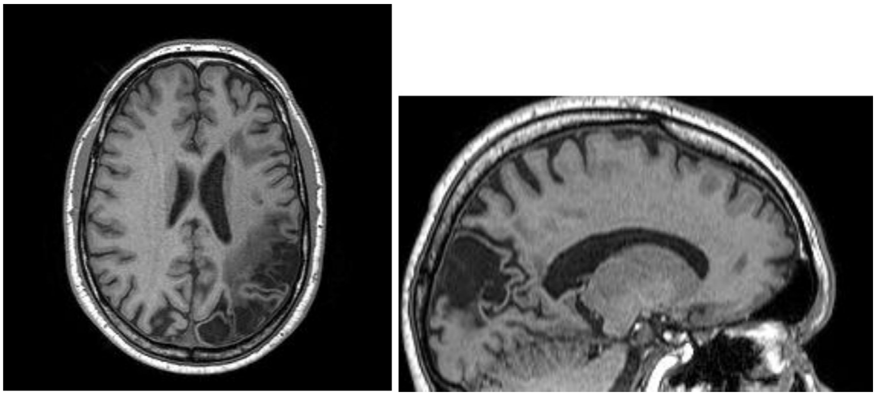


Fig. 3 – Radiological CT Scan for case CT.

Lesser & Coltheart, 1992) battery. Specifically, we selected subtests that examined word and non-word repetition, rhyme judgement, homophone judgement, visual and auditory lexical decision, picture naming and reading aloud. In addition we used Pyramids and Palm Trees (Howard & Patterson, 1992) as a measure of semantic performance. This test presents the participant with 3 images (or words) simultaneously. The stimulus at the top of the triangle is the reference image. One of the two items below represents a semantically-related concept and the participant has to determine which it is. As well as the non-word reading task from the PALPA battery, we also presented CT with the 100 non-words selected by Weekes (1997).

CT was presented with lists of single-letter stimuli for immediate serial recall. The number of items in each list was determined relative to CT's pSTM span for visual and auditory information. Phonologically similar lists were comprised of the letters B, C, D, E, P, T and V. For the phonologically dissimilar lists, the possible letters were F, J, Q, R, S, W and Y. In the auditory presentation, lists were 4 items long and letters were spoken by the experimenter at a rate of 1 per second. In the visual modality, the lists were 7 items long and the letters were presented on paper at the same rate as for auditory presentation. CT was tested for immediate serial recall of these lists over 5 sessions, each session containing 10 phonologically similar and 10 phonologically dissimilar lists. The final part of the study presented CT with lists of 7 letters for immediate serial recall, inspired by the work of Best and Howard (2005). For visually similar lists, the available letters were H, M, N, U, V, W, and Y. The visually dissimilar lists contained the letters A, I, J, L, P, X, and Z. There were 10 visually similar and 10 visually dissimilar lists – with presentation of visually similar or dissimilar items presented in blocks. These lists were presented twice in separate sessions, and only in written form.

4.3. *Control participants*

A total of 18 age-matched controls (10 males, 8 females – aged 59–70) were recruited for the immediate serial recall task that manipulated *phonological* similarity (11 for the written version, 7 for the auditory presentation). From this group, a sub-group of 12 age-matched control participants (5 males, 7 females – aged 59–70) were recruited for the immediate serial recall task that manipulated *visual* similarity. The control participants were all native speakers of English without history of stroke, dementia or other brain injury. CT's performance on the standardised test battery was compared to the normative data published with the tests, unless otherwise stated in the results section.

5. Results

5.1. *Neuropsychological data*

We administered a large number of standardised tests to CT to provide an initial assessment of his cognitive function. Table 1 presents accuracy data for CT in each of the tests administered to him, alongside control data (drawn from the test manuals unless otherwise noted). Some of the tasks we presented to CT are particularly relevant to a) our conclusion that he has conduction aphasia and b) our subsequent investigation of his pSTM performance, hence we have described these tests and results in greater detail below.

5.2. *Speech perception*

We assessed CT's ability to perceive speech at both the phonological and the semantic level. For the phonological level, we used word and non-word minimal pairs discrimination tasks – subtests 1 and 2 from the PALPA battery. In each of these tasks the patient was presented with 72 pairs of stimuli, with each item spoken by the experimenter approximately 1 sec apart. In half of trials the two members of the pair are identical; in half the two members differ by one phonological feature (voice, manner or place of articulation). Differences between pairs may occur at the initial or final positions, or the pairs may be metathetically related (i.e., the order of the sounds are reversed). All stimuli are monosyllabic with a CVC structure. Poor performance on these tasks is likely to indicate that the ability to parse and decode auditory phonological input is compromised. The administration and comparison of the word and non-word versions allows for an assessment of whether lexical information can be used to reduce any deficit. CT performed as well as controls in both versions ($p > .05$) as

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

Table 1 – Basic neuropsychological test data. Control data taken is from the original published measures unless otherwise indicated, SD in parentheses where available.

	CT	Controls
General Cognitive skill		
Mini-Mental State	28/30	29/30
Ravens Progressive Matrices	9/12	10/12
Rey Figure Copy	36/36	35/36
Auditory Digit Span	5 forwards: 3 backwards	
Visual Digit Span	6 forwards: 5 backwards	
Visual Processing		
BORB Minimal Feature	25/25	23/25
BORB Foreshortened	25/25	22/25
BORB Object Decision	124/128	115/128
VOSP – Shape Detection	19/20	20/20
VOSP – Position Discrimination	19/20	20/20
Semantics		
PPT pictures	50/52	50/52
PPT written	52/52	50/52
PALPA 50 Auditory Synonyms	58/60	None available
PALPA 49 Visual Synonyms	60/60	None available
ADA Word – Picture matching	65/66	Cut-off 63/66
Repetition		
PALPA 9 High Imageability, High Frequency words	19/20	19.81/20 (.60)
PALPA 9 High Imageability, Low Frequency words	18/20	19.52/20 (.93)
PALPA 9 Low Imageability, High frequency words	15/20	19.81/20 (.60)
PALPA 9 Low Imageability, Low Frequency words	13/20	19.67/20 (.58)
PALPA 9 Non-words	2/80	75.94/80 (6.72)
Phonological processing		
PALPA 28 Homophone judgement	60/60	54.81/60 (2.77)[a]
PALPA 15 Written rhyme judgement SPR	15/15	54.20/60 (2.63)[b]
PALPA 15 Written rhyme judgement SPC	14/15	
PALPA 15 Written rhyme judgement PR	15/15	
PALPA 15 Written rhyme judgement PC	14/15	
PALPA 15 Auditory rhyme judgement SPR	14/15	55.65/60 (2.48)[b]
PALPA 15 Auditory rhyme judgement SPC	14/15	
PALPA 15 Auditory rhyme judgement PR	14/15	
PALPA 15 Auditory rhyme judgement PC	15/15	
PALPA 2 Minimal Pairs (words)	70/72	
PALPA 1 Minimal Pairs (non-words)	68/72	
Lexical decision		
PALPA 25 (Visual) High Imageability, High Frequency words	15/15	14.79/15 (.51)
PALPA 25 (Visual) High Imageability, Low Frequency words	15/15	14.58/15 (.58)
PALPA 25 (Visual) Low Imageability, High frequency words	15/15	14.92/15 (.41)
PALPA 25 (Visual) Low Imageability, Low Frequency words	15/15	14.71/15 (.75)
PALPA 25 (Visual) Non-words	60/60	59.88/60 (.45)
PALPA 5 Auditory lexical decision		
Reading		
PALPA 31 Words	79/80	79.4/80 (.80)
PALPA 36 Non-words 3 letters	6/6	5.77/6 (.71)
PALPA 36 Non-words 4 letters	6/6	5.89/6 (.43)
PALPA 36 Non-words 5 letters	6/6	5.57/6 (.90)
PALPA 36 Non-words 6 letters	6/6	5.65/6 (.85)
Weekes (1997) Low frequency	100/100	99/100
Weekes (1997) High frequency	100/100	99/100
Weekes (1997) Non-words	96/100	97/100
Picture naming		
PALPA 53	40/40	39.80/40 (.35)
Graded Naming Test	20/30	20.4/30 (4.1)[c]

Note: BORB = Birmingham Object Recognition Battery (Riddoch & Humphreys, 1993), VOSP = Visual Object and Space Perception Battery (Warrington & James, 1991), PPT = Pyramids and Palm Trees (Howard & Patterson, 1992), PALPA = Psycholinguistic Assessment of Language Performance in Aphasia (Kay et al., 1992). SPR = spelling pattern rhyme, SPC = spelling pattern control, PR = phonological rhyme, PC = phonological control.

[a] Control data from 21 young adult participants, taken from Nickels and Cole-Virtue (2004).

[b] Control data from 20 typically ageing adults, taken from Harley et al. (2013). The original paper does not split data into conditions.

[c] Control data taken from Warrington (1997). Weekes (1997) did not present SD for his participants (accuracy has been extrapolated from his Fig. 1).

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

determined by Crawford's *t*-tests (Crawford & Howell, 1998). His accuracy on the non-word discrimination task (94%) was slightly lower than for the word task (97%), but neither fell outside of the normal range. Therefore, CT appears to have intact phonological decoding ability.

For the sematic level of speech perception, we administered the spoken word to picture matching task from the ADA Comprehension Battery (Franklin, Turner & Ellis, 1992). In this task, the patient is presented with a series of 66 trials in which they must point to the picture that represents an auditorily presented word from an array of 4 line drawings. On each trial there are two unrelated pictures alongside the target, plus a third distractor that is related to the target in phonology, semantics, or both. The manual for this test indicates that healthy controls should make no more than 3 total errors on this task. CT scored 65 correct out of a possible 66, hence clearly performs in the normal range. This indicates that CT has no problems with comprehension of auditory input and that any pattern of errors in subsequent testing is unlikely to be the result of difficulty in word comprehension. This is further supported by CT's performance on semantic tests (see Pyramids & Palm Trees test and synonym judgement in Table 1).

5.3. *Speech production*

Although CT's spontaneous speech was generally fluent, we also formally assessed CT's speech production using two picture-naming tasks (PALPA 53 and the Graded Naming Test, McKenna & Warrington, 1983). The PALPA task presents 40 pictures of common objects, and the patient is required to speak the appropriate noun to describe the object. In this, CT scored 40 out of 40. The Graded Naming Test asks participants to provide the appropriate noun for 30 pictured objects which become increasingly obscure and difficult as the test progresses. In a revalidation of the Graded Naming Test, Warrington (1997) reported that healthy control performance was, on average, 20.4 out of 30 (SD = 4.1). CT scored 20 out of 30 on this task, again performing in the normal range. In sum, the evidence suggests that CT has no particular deficit in speech production. Therefore we can assume that any deficits in performance for the tasks described later in this paper are not attributable to articulatory speech problems.

5.4. *Repetition and reading*

The cardinal feature of conduction aphasia is a marked reduction in the ability to repeat an auditorily presented item. To test this, we again turned to the PALPA battery. PALPA 9 presents 160 items for immediate repetition, split into equal numbers of real words and non-words. The word list varies frequency and imageability orthogonally, with 20 items in each cell. Theoretically, as we alluded to in the introduction, any of the items (irrespective of lexicality) could be repeated successfully without recourse to lexical or semantic representations. For nonword repetition, of course, there are no lexico-semantic representations to access in any case, so a direct repetition route is the only viable option. Words could be repeated lexically, with or without semantic information becoming involved – repetition via meaning would be likely to result in better performance for high versus low imageability words. CT's repetition performance can be characterised as follows: a clear advantage for words over nonwords, an imageability effect, and no frequency effect. CT correctly repeated only 2 out of 80 nonwords correctly (control mean = 75.94, SD = 6.72). We note two points concerning non-word repetition in our case. Firstly, other authors (e.g., Jacquemot et al., 2007) have reported significant effects of syllable or phoneme length in non-word repetition in conduction aphasia cases. We did not observe such an effect in CT, but this may be due to his remarkably poor overall performance – CT simply could not repeat non-words correctly whatever length they were. The second observation is that 30 out of the 78 errors CT made were lexical captures, which hints at an attempt to use a lexical repetition route even when the stimuli were inappropriate for such a strategy. His word repetition was also impaired versus controls [65/80 for CT, mean = 78.81/80, SD = .69 for controls; Crawford's t (20) = 19.554, $p < .001$]. This lexicality effect in repetition is characteristic of conduction aphasia. Of CT's 15 errors in the word repetition task, 12 came in low imageability trials which constitutes a significant effect [χ^2 (1) = 5.251, $p < .05$]. Accuracy was statistically similar across high and low frequency items [6 *vs* 9 errors respectively, χ^2 (1) < 1]. This suggests that CT was making use of semantic information in repetition and that this was defective in some way, but the *types* of errors that CT made does not necessarily bear this out. He committed 6 formal errors (clue-blue), offered 5 neologisms (folly-forrow), and gave no response to 4 items but did not make any semantic errors at all. It is also worth noting that there were no obvious syllable length effects in word repetition (errors were for 3 one syllable words, 6 two syllable, 5 three syllable and 1 four syllable). We acknowledge that PALPA 9 does not explicitly vary or control syllable length, but it is nevertheless interesting in that syllable length effects in word repetition may indicate a deficit in the phonological output buffer.

Subtest 31 of the PALPA battery presents the same word items from the repetition task described above for reading aloud. By testing reading and repetition with the same items it is possible to check whether the repetition deficit is attributable to a production deficit – the production deficit would affect repetition and reading similarly. CT, however, successfully read 79 of the 80 words. We also presented CT with non-word stimuli for reading aloud using PALPA 36. This subtest presents 24 non-words (6 each of 3, 4, 5, and 6 letters long). CT performed at ceiling for this task. Finally, we presented the stimuli from Weekes' (1997) study, that manipulated length and lexicality. In this set there are 100 low frequency words, 100 high frequency words and 100 non-words, with equal numbers of 3, 4, 5 and 6 letter items. CT correctly read all of the words correctly and made only 5 errors on the non-words, which is comparable to the healthy participants in Weekes' paper. Again, there was no observable length effect (2 errors for 3 letter words, 2 for 4 letter words and 1 for 6 letter words). Taken together, the results from the reading aloud tasks indicate that CT's repetition deficit is unlikely to have been caused by a difficulty in producing the necessary phonemes, and instead reflects a specific deficit in converting aural input to oral output.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

5.5. *Rhyme and homophone judgement*

Following the logic of Nickels et al. (1997), we administered rhyme and homophone judgement tasks in order to attempt to pinpoint where in pSTM any deficit might be. We used PALPA 16 (rhyme judgement) and PALPA 28 (homophone judgement) as described below. To pre-empt what follows, CT showed no deficits in these tasks. The rhyme judgement task from the PALPA battery contains 60 pairs of words, half of which rhyme. Pairs belong to four categories of 15 pairs each. Rhyming pairs can share orthography and phonology (spelling pattern rhymes, like *match* and *hatch*) or phonology only (phonological rhymes, like *you* and *two*). Non-rhyming pairs can also share orthography (spelling pattern controls, like *five* and *give*) or not (phonological controls, like *sort* and *part*). The participant simply has to indicate whether the pair rhymes or not. Harley, Oliver, Jessiman and MacAndrew (2013) administered these tests with a group of 20 typically ageing older adults and reported that mean accuracy for the written version was 54.20 out of 60 (SD = 2.63), while the mean for the auditory version was 55.65 (SD = 2.48). We used these data as a normative estimate for comparison with CT, who also performed written and auditory versions of the rhyme judgement task. CT scored in the normal range as determined by Crawford's *t*-tests [58/60 for written, t (19) = 1.410, *p* > .1; 56/60 for auditory, t (19) < 1]. The homophone judgement task also contains 60 pairs. In this task, the participant has to determine whether the two members of the pair sound *exactly* alike. There are 10 pairs of homophonic regular words (e.g., *sea* and *see*), 10 homophonic pairs in which one or both words are irregular (e.g., *quay* and *key*) and 10 pairs of homophonic non-words (e.g., *zole* and *zoal*). This allows for graphene-phoneme assembly to be assessed in the non-word trials and for lexical reading to be assessed using the irregular word pairs. The remaining 30 trials (20 word pairs, 10 non-word pairs) contain non-homophonic items that are as visually similar as the members of the homophonic pairs. CT was correct on every trial.

In sum, the performance of CT on the rhyme and homophone judgement tasks can be accommodated in a dual buffer model of pSTM such as Monsell (1987), by assuming that a) the phonological input buffer is intact and can be used to complete auditory rhyme judgement, b) the phonological output buffer is intact and can be used to complete homophone decision and c) information can be passed from the output to the input buffer to allow correct written rhyme judgements to be made. When we consider the other tasks that have been described so far it appears that CT has an impairment in converting phonological input to phonological output, particularly if lexical or semantic knowledge cannot be used to help. Specifically, we argue that his performance can be accommodated parsimoniously by positing that he has a dysfunctional link from the input buffer to the output buffer in pSTM – or more simply, that CT has only half a phonological short-term memory rehearsal loop.

5.6. *Immediate serial recall - phonological similarity*

CT was presented with lists of 4 letters for auditory serial recall. The length of the list was constrained by his auditory digit span. CT's recall of the list items were scored as correct only if he reported the presented letter in the list position that it had been presented. This meant that CT was required to offer the same number of items in a sequence, but in instances when he could not recall the letter he would say "pass" (e.g., for sequence E, C, T, P, B, V, D, he might say "E, C, T, pass, B, D, V"). Overall, CT performed better for auditory (75% correct) versus visual (68% correct) presentation [χ^2 (1) = 6.14, *p* < .01]. We note that this is not the usual pattern observed in cases of pSTM impairment but in this instance we attribute this finding to the disparity in the number of items in the visual versus auditory lists - each correctly recalled letter constitutes a greater proportion of the list in the auditory presentation. In the auditory version of the task, CT showed a significant phonological similarity effect [χ^2 (1) = 15.11, *p* < .001]. The same was not true in the visual presentation, where CT's recall performance was not affected by phonological similarity [χ^2 (1) < 1]. These patterns are explored in more detail below. The control participants, however, *did* show a phonological similarity effect in the visual presentation such that recall was significantly poorer for the similar (55%) than dissimilar (80%) lists [t (10) = 7.80, *p* < .001]. As the auditory presentation was only 4 items per list, the control participants scored at or near ceiling in phonologically similar and phonologically dissimilar lists. As a result, no phonological similarity effect was observed.

Table 2 shows the proportion of letters CT correctly recalled in each position across the test sessions for the auditory presentation. The mean performance for our control group is also included.

CT's overall accuracy was significantly lower that control participants for both phonologically similar [t (1) = 11.81, *p* < .001] and phonologically dissimilar [t (1) = 6.29, *p* < .001] lists. Crawford's *t*-tests also indicated that CT's recall was significantly lower than the control group for phonologically similar letters in the second and third positions in the list, and for letters in the penultimate and final positions in the phonologically dissimilar condition (all *p* < .001). It should be noted here that Crawford's *t*-test compares the score of a single case to a *distribution* of normative responses to determine whether the patient is performing outside of what could be expected for a non-impaired participant - hence the test cannot be applied in instances where the normative population has a standard deviation of zero. For this reason it was not possible to formally assess CT's performance in any list position for which our control participants scored 100%. That said, one can assume that *any* error under circumstances where a normative sample has performed at ceiling represents a significant impairment (i.e., he was impaired at the final list position in the similar lists too). We entered the number of letters CT correctly recalled at each list position into separate Friedman's ANOVAs for phonologically similar and phonologically dissimilar lists. For the phonologically similar lists, the effect of list position was significant [χ^2 (3) = 12.894, *p* < .001], and it was also significant in the phonologically dissimilar lists [χ^2 (3) = 11.769, *p* < .01]. In neither list did any of the pairwise comparisons (Wilcoxon signed ranks tests, Bonferroni correction applied) reach significance. A 2 (similarity) x 4 (list position) repeated measures ANOVA for the control participant revealed no significant main effects or interaction term, probably because the participants were at or near

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

Table 2 – Mean number of letters recalled correctly in each list position for CT, alongside mean performance for the control group (SD in parentheses, where available). There were no significant similarity effects at any list position (Bonferroni corrected *t*-tests).

	CT			Controls		
	Similar	Dissimilar	Difference (Dissimilar −Similar)	Similar	Dissimilar	Difference (Dissimilar −Similar)
Position 1	8	10	2	10	10	0
Position 2	7.8	10	2.2	9.7 (.49)	10	.3
Position 3	6.2	8.2	2	9.4 (.79)	9.9 (.38)	.5
Position 4	3	6.8	3.8	10	9.9 (.38)	−.1

ceiling. A final observation relates to CT's performance in recall of items in the final position of to be recalled sequences – as mentioned earlier, in previous work with healthy participants, the auditory modality shows the biggest advantage at the recency position (i.e., final items) of the serial position curve (Conrad & Hull, 1968). For CT (like many impaired pSTM cases) this recency advantage is not present.

We initially presented CT with lists of 4 letters long for visual immediate serial recall as well, but he performed at ceiling. We also presented lists of 5, and then of 6 letters in length. For both of these, the overall accuracy was high (91% in 5 letter lists, 86% in 6 letter lists), so we ultimately presented lists containing 7 letters for the examination of visual immediate serial recall so that any patterns of performance were not likely to be confounded by ceiling effects. Fig. 4 shows the proportion of letters CT correctly recalled in each position across the test sessions for the visual presentation. The mean performance for our control group is also included.

As a first step in the analysis of the effect of phonological similarity on recall for visually presented lists, we entered the number of letters correctly recalled by our controls into a 7 (list position) x 2 (similar *vs* dissimilar) repeated measures ANOVA. This revealed a significant main effect of phonological similarity [F (1, 10) $= 60.853$, $p < .001$, $\eta p^2 = .589$].

With the advantage being for dissimilar lists. There was also a main effect of list position [F (6, 60) $= 33.421$, $p < .001$, $\eta p^2 = .770$]. Post hoc Bonferroni corrected *t*-tests revealed that letters in position 1 were recalled significantly more often than letters at positions 3 onwards, that letters in position 2 were recalled significantly more often than letters in position 4 onwards, and that letters in position 3 were recalled significantly more often than letters in position 5 onwards (all $p < .05$). Overall recall was statistically similar at positions 4, 5, 6 and 7 (ps > .1). The interaction between similarity and list position was also significant [F (6, 60) $= 17.674$, $p < .001$, $\eta p^2 = .643$], with greater recall differences between similar and dissimilar lists at later list positions (see Fig. 4 and Table 3).

Although we could not formally assess whether there was a similarity × list position interaction in CT's performance, we conducted Friedman's ANOVAs separately for the phonologically similar and phonologically dissimilar lists. In phonologically similar lists, there was a significant effect of list position [χ^2 (6) $= 28.260$, $p < .001$]. A series of Wilcoxon tests were conducted to further interrogate this effect, but none of the pairwise comparisons survived the Bonferroni correction. The same was true in the analysis of phonologically dissimilar lists – the overall effect of list position was significant [χ^2 (6) $= 28.143$, $p < .001$] but pairwise comparisons did not reach significance once the Bonferroni correction was applied. Table 3 shows that there was no significant similarity effect at any list position.

Table 3 – Mean number of letters correctly recalled at each list position, in each list, by CT and controls (SD in parentheses for controls). Significant differences between phonologically similar and dissimilar list performance (t-tests for controls, Wilcoxon signed ranks for CT) at each list position have been indicated.

CT	Controls		
Difference (Dissimilar – Similar)	Similar	Dissimilar	Difference (Dissimilar – Similar)
0	9.55 (.69)	9.73 (.47)	.18
.8	8.18 (1.25)	9.27 (1.19)	1.09
0	6.82 (1.47)	8.55 (1.13)	1.73*
−.6	5.36 (1.50)	8.18 (1.17)	2.82*
−2.2	3.36 (1.63)	6.18 (2.71)	8.82*
−.8	3.00 (1.55)	6.09 (2.70)	3.09*
.6	2.73 (1.42)	8.18 (1.40)	5.45*

*$p < .05$, **$p < .01$ (Bonferroni corrected).

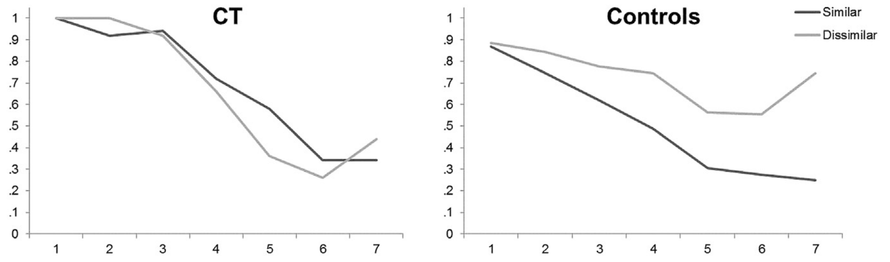


Fig. 4 – Proportion of letters recalled correctly in each list position, in phonologically similar and dissimilar lists, for CT and controls.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

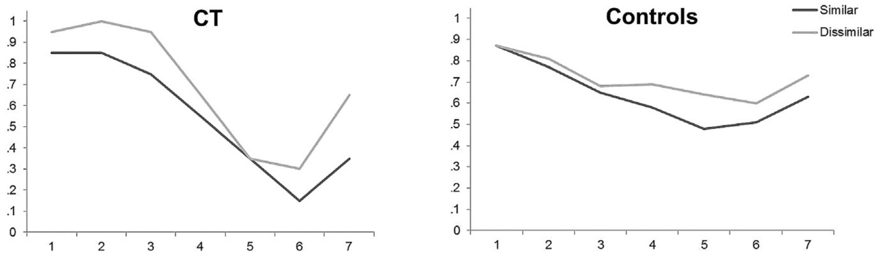


Fig. 5 – Proportion of letters recalled correctly in each list position.

We compared CT's performance to that of the control group for phonologically similar, and phonologically dissimilar, lists using Crawford's t-tests (Crawford & Howell, 1998) both in overall proportion correct, and split by list position. Overall, the proportion of letters remembered correctly by CT was significantly greater than the control group in phonologically similar lists [$t\ (1) = 4.89$, $p < .001$] and significantly lower in phonologically dissimilar lists [$t\ (1) = 4.73$, $p < .001$]. CT also performed within the normal range in all 7 list positions (all $p > .1$) for the phonologically similar lists. In fact, only the proportion of phonologically dissimilar letters correctly recalled in the final list position was significantly different from the control participants [$t\ (1) = 2.92$, $p < .05$] with CT scoring lower. Thus overall, CT has a higher Visual ISR span than Auditory ISR span, but only shows phonological similarity effects for auditory but not visual ISR.[1] Thus in all cases there is a dissociative pattern of performance across VISR and AISR tasks for CT.

In summary for AISR testing, CT, a) shows an overall lower performance versus controls, b) shows an overall phonological similarity effect and c) shows no evidence of a recency effect. This contrasts markedly for VISR performance, in which case, (a) performance is much improved relative to AISR, (b) no evidence of a phonological similarity effect and (c) an upturn in recall at the final list position (though not significant). Indeed, with respect to (b), in the visual modality, CT's recall of phonologically similar items is significantly better than controls. Overall, we argue that this pattern indicates that CT is not converting visual stimuli into a phonological code, and likely is using visual information to support his performance (in line with the findings of Vallar et al. (1997) and Best and Howard (2005)). In order to investigate this issue further we examined whether CT's performance might be impacted by *visual similarity* effects in VISR.

5.7. *Immediate serial recall - visual similarity*

CT's recall for lists of visually similar and visually dissimilar lists was scored in the same way as for the phonological manipulations in the previous section. These proportions, along with the mean performance of our control group, are presented in Fig. 5. Chi square comparison between CT's overall accuracy in the visually similar (55%) and visually dissimilar (69%) lists indicated that his performance was significantly different, in the predicted direction [X^2 (1) = 4.178, $p < .05$]. A significant difference in performance was also observed in our control participants as a group [$t\ (11) = 3.60$, $p < .05$] such that accuracy for visually similar letters (64%) was lower than for visually distinct letters (72%) – though only 2 of the controls actually showed a significant visual similarity effect when analysed individually. Crawford's t-tests determined that CT did not score outside of the normal range for either list overall, nor for his correct recall of letters presented in any list position (all $p > .1$).

We assessed the effects of similarity and list position on our control participants' recall by conducting a 2 × 7 repeated measures ANOVA. This revealed an overall main effect of similarity [$F\ (1, 11) = 11.943$, $p < .01$, $\eta p^2 = .521$], such that recall was better for visually dissimilar letters. There was also an overall main effect of list position [$F\ (6, 66) = 14.018$, $p < .001$, $\eta p^2 = .560$]. Post hoc tests were performed, but only the significant differences (all $p < .05$, Bonferroni corrected) will be reported for brevity. It was identified that recall was significantly better at position 1 than for positions 3, 5, 6, and 7. Letters at position 2 were recalled better than for positions 5 and 6. Recall for letters presented at position 3 and 4 was better than for letters presented at position 5. In summary, then, the control participants performed significantly better at earlier list positions. Although there was a numerical increase in recall at the penultimate and final list positions, this did not constitute a statistically significant recency effect. The interaction was not significant ($p > .05$). Again, we performed separate Friedman's ANOVAs on the number of letters recalled by CT in the similar and dissimilar lists. In both cases, there were significant list position effects [visually similar letters χ^2 (6) = 11.450, $p < .01$; visually dissimilar letters χ^2 (6) = 10.598, $p < .05$]. None of the pairwise comparisons in either list reached significance (all $p > .1$). Table 4 presents the number of letters recalled by CT and controls at each list position of the similar and dissimilar lists. Though none of the comparisons survived Bonferroni correction, there are some clear numerical differences between similar and dissimilar list recall. For example, CT recalled twice as many letters in the penultimate position of the dissimilar lists than he did at the same position in the similar lists.

In sum, we predicted that CT would be affected by the visual similarity manipulation, and that is indeed what appears to be present overall. As a consequence we would argue that as a consequence of his pSTM impairment, CT has shifted his emphasis to visual information when available in that modality – a pattern we suggest is similar to other published reports (Vallar et al. (1997) and Best and Howard (2005)).

---

[1] We also think it is worth noting that CT also had an overall accuracy of 68% when we presented him with lists of 4 *words*, which is a more difficult task. Here, too, CT was considerably impaired relative to controls (who all performed at or near ceiling), but he again showed no phonological similarity effect.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

Table 4 – Mean number of letters correctly recalled at each list position, in each list, by CT and controls (SD in parentheses for controls). Significant differences between visually similar and dissimilar list performance (t-tests for controls, Wilcoxon signed ranks for CT) at each list position have been indicated.

Position	CT			Controls		
	Similar	Dissimilar	Difference (Dissimilar – Similar)	Similar	Dissimilar	Difference (Dissimilar – Similar)
1	8.5	9.5	1.0	8.67 (1.07)	8.83 (1.40)	.17
2	8.5	10	1.5	7.50 (1.83)	7.92 (1.44)	.42
3	7.5	9.5	2.0	6.33 (1.78)	6.67 (2.39)	.33
4	5.5	6.5	1.0	5.67 (2.99)	6.92 (1.98)	1.25
5	4.0	3.5	−0.5	4.17 (2.25)	6.17 (2.41)	2.00*
6	1.5	3.00	1.5	4.75 (2.60)	5.50 (2.65)	.75
7	3.5	6.5	3.0	6.00 (2.76)	7.00 (1.81)	1.00

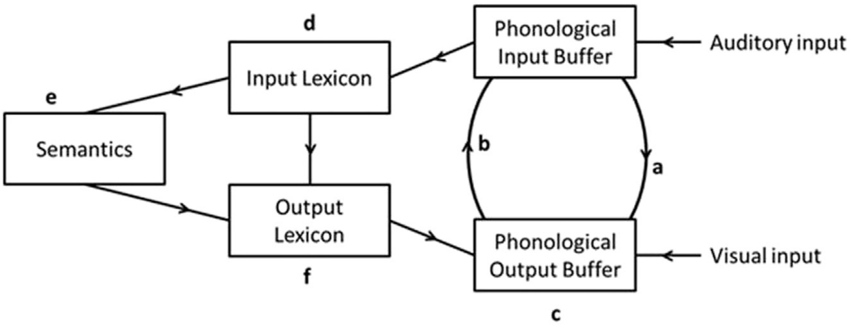
*$p < .05$, **$p < .01$ (Bonferroni corrected).

6. Discussion

Our main aim in this paper was to explore the functioning of pSTM across a variety of phonological tasks in a case of conduction aphasia, and interpret our findings with respect to a key dual buffer model suggested by Monsell (1987). We also sought to examine CT's performance across visual and auditory versions of ISR tasks and to determine if there may be any evidence of a visual similarity effect for our case. The findings can be summarised as follows. First, our conduction aphasic could perform nonword reading, lexical decision, semantic association, homophone judgement, written and spoken rhyme judgement tasks with the same level of accuracy as non-impaired controls, in spite of his reduced short term memory span and severe deficit in non-word repetition. Second, CT showed phonological similarity effects in auditory, but not visual, immediate serial recall tasks, despite superior performance on the latter relative to the former. Finally, our investigation of VISR letter visual similarity effects for both CT and controls showed that both were affected by the level of visual similarity between the letters in the lists. We consider each of these key patterns of performance in greater detail below, but conclude all can be accounted for under Monsell's (1987) framework see Fig. 6 below (in line with the account of Nickels et al., 1997).

6.1. *CT – an impairment of half the phonological ‘loop’*

As we have established, CT's pattern of impairment is largely confined to phonological tasks that involve mapping spoken input to output (his nonword repetition in particular is severely disrupted). This is despite striking good performance at a great many visual phonological tasks (rhyme judgement/homophone judgement) and auditory phonological tasks (rhyme judgement/lexical decision). As a consequence, with reference to Fig. 6 and the baseline data presented previously in Table 1, we interpret CT's impairment as resulting in severe disruption to the link from the phonological input buffer to the phonological output buffer – much as Jacquemot et al. (2007)



Damage to model	Homophone judgement	Auditory lexical decision	Written rhyme judgement	Auditory rhyme judgement
Input buffer		x	x	x
Output buffer	x	x	x	
Input-to-output link				
Output-to-input link		x	x	
CT Performance	100%	96%	97%	93%

Note: x indicates that an impairment would be predicted

Fig. 6 – Dual-Buffer pSTM model and predictions about task impairments to functional damage. CT's performance is drawn from the baseline neuropsychological testing undertaken and reported in Table 1.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

explained the performance of FA. In theory, a disconnection in the conversion from input to output would severely reduce short term memory span by preventing the use of a rehearsal loop. However, if the input and output buffers themselves were intact, tasks that could be performed using either of these components would remain unimpaired. This matches the pattern we observed in CT.

In line with the narrative suggested by Nickels et al. (1997), auditory rhyme judgements can be made solely in the phonological input buffer; homophone judgements can be made solely in the phonological output buffer. CT completed both of these tasks with normal levels of accuracy. Visual rhyme judgements require that written input is converted to a phonological code and passed from output buffer to input buffer before a decision can be made. In CT, visual rhyme judgement was in the normal range, as might be expected if the output-input link had not been damaged. We note that FA (Jacquemot et al., 2007) had an impairment in written rhyme judgement in spite of the fact that the authors of that paper argued that FA's output-input link was intact. However, the pattern of errors FA committed in this task parallelled her performance in reading aloud, and Jacquemot et al. (2007) attributed, therefore, the rhyme judgement deficits to reading errors rather than pSTM processes. Non-word repetition, on the other hand, relies on direct connections from the phonological input buffer to the phonological output buffer – CT was severely impaired at non-word repetition, indicating a problem with the input-output link. We argue that his preserved ability to read and recognise written words and perform semantic tasks indicates that his deficit is limited to pSTM (in contrast to FA, Jacquemot et al., 2007).

Thus, overall, the reported pattern of performance matches the predictions of Monsell (1987) and Nickels et al. (1997) and can be readily accommodated in a model that proposes separate phonological input and output buffers, and distinct conversion processes in each direction. Jacquemot et al. (2007) reported a conduction aphasic case that is strikingly similar to our patient, CT. In their discussion, Jacquemot et al. noted that it would be potentially possible for rhyme judgement tasks to be completed on the basis of *orthography* rather than phonology, by comparing the visual representations of the items in the pairs (either as presented in the written version, or as computed in the auditory version), although they discarded this possibility in FA. We do not think that orthography is at the root of CT's preserved rhyme judgement performance either. CT's performance was near ceiling irrespective of whether the rhyming pairs shared spelling patterns.

6.2. *CT – the consequences of his disruption on ISR tasks*

As we established earlier, we argue that CT's impairment reflects damage to the input-output pathway (or half of the phonological ‘loop’). At this point our discussion moves to issues relating to immediate serial recall in both the auditory and the visual domain. As stated in the introduction, it is common for healthy participants to perform better at auditory than visual serial recall, whereas the pattern is reversed in patients with pSTM deficits (Best & Howard, 2005; Howard, 1995; Shallice & Vallar, 1990; Warrington & Shallice, 1969; 1972). CT, too, showed better levels of accuracy in visual serial recall than in auditory presentation. Of greater interest, though, is the varying susceptibility of CT's serial recall to phonological and visual similarity between items. The phonological similarity effect is observed in auditory or written presentation in normal participants (e.g., Besner & Davelaar, 1982) except when under conditions of articulatory suppression, whereupon the phonological similarity effect is abolished for written presentation only (e.g., Vallar & Baddeley, 1984). Given that the evidence suggests that CT has a damaged rehearsal mechanism in pSTM, we suggested that it would be as if he was performing *any* short-term memory task under articulatory suppression. This prediction was borne out in the data – CT showed significant phonological similarity effects in auditory, but not visual, serial recall (in fact, there was some evidence of a reverse effect – better recall of phonologically similar letters – relative to controls in this case).

In sum, CT's pattern of performance is consistent with other patients with ISR impairments such as Vallar et al. (1997) and Best and Howard (2005). Finally, it appears that CT's immediate serial recall performance for visually similar letters (presented in written form) is worse than for visually dissimilar items (in effect a corollary of the phonological similarity effect) – again suggesting that for CT, given his phonological rehearsal impairments, he is to some extent using some form of ‘*visual information*’ as shift of emphasis to achieve better performance on VISR (in line with the suggestions of Baddeley, 1986). Intriguingly, there is other evidence that under the context of articulatory suppression, some similar resource or encoding shift occurs for normal participants (see, Tree, Longmore, Majerus, et al., 2011; Tree, Longmore, & Besner, 2011). This finding is also in line with work by Best and Howard (2005) who reported similar visual letter confusability effects for their patient, who even spontaneously reported “*The W and M are mixing me up*”. In all then, there is increasing evidence to suggest that in VISR tasks there is likely the potential for shifting of encoding and storage emphasis from phonological to visual information depending on the task demands (or the nature of patient impairment). Further work is needed to explore this in greater detail.

REFERENCES

Allport, D. A. (1984). Auditory-verbal short-term memory and aphasia. In H. Bouma, & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes*. London: Erlbaum.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

Bartha, L., & Benke, T. (2003). Acute conduction aphasia: An analysis of 20 cases. *Brain and Language, 85*, 93–108.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology, 36*(4), 701–711.

Best, W., & Howard, D. (2005). “The W and M are mixing me up”: Use of a visual code in short-term memory tasks. *Brain and Cognition, 58*, 274–285.

Butterworth, B. (1992). Disorders of phonological encoding. *Cognition, 42*, 261–286.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008

Code, C., Tree, J. J., & Dawe, K. (2009). Opportunities to say 'yes': Rare speech automatisms in a case of progressive nonfluent aphasia and apraxia. *Neurocase, 15*(6), 445–458.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429–432.

Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science, 10*(4), 135–136.

Copeland, D. E., & Radvansky, G. A. (2001). Phonological similarity in working memory. *Memory & Cognition, 29*, 774–776.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*, 482–486.

Franklin, S., Buerk, F., & Howard, D. (2002). Generalised improvement in speech production for a subject with reproduction conduction aphasia. *Aphasiology, 16*, 1087–1114.

Franklin, S., Turner, J. E., & Ellis, A. W. (1992). *ADA comprehension battery: Action for dysphasic adults*. London: Action for Dysphasic Adults.

Goodglass, H. (1992). Diagnosis of conduction aphasia. In S. E. Kohn (Ed.), *Conduction aphasia*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gvion, A., & Friedmann, N. (2012). Phonological short-term memory in conduction aphasia. *Aphasiology, 26*(3–4), 579–614.

Harley, T., Oliver, T. M., Jessiman, L. J., & MacAndrew, S. B. G. (2013). Ageing makes us dyslexic. *Aphasiology, 27*, 490–505.

Howard, D. (1995). Short term recall without short term memory. In R. Campbell, & M. A. Conway (Eds.), *Broken memories: Case studies in memory impairment*. Oxford: Basil Blackwell.

Howard, D., & Franklin, S. (1988). *Missing the meaning?* Cambridge, MA: MIT Press.

Howard, D., & Nickels, L. (2005). Separating input and output phonology: Semantic, phonological, and orthographic effects in short-term memory impairment. *Cognitive Neuropsychology, 22*, 42–77.

Howard, D., & Patterson, K. (1992). *Pyramids and palm trees: A test of semantic access from pictures and words*. Bury St. Edmunds, England: Thames Valley Test Company.

Jacquemot, C., Dupoux, E., & Bachoud-Levi, A. (2007). Breaking the mirror: Asymmetrical disconnection between the phonological input and output codes. *Cognitive Neuropsychology, 24*, 3–22.

Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language processing in aphasia*. Hove: Lawrence Erlbaum Associates.

Kohn, S. E. (1992). Conclusions: Toward a working definition of conduction aphasia. In S. E. Kohn (Ed.), *Conduction aphasia* (pp. 151–156). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Kohn, S. E., & Smith, K. L. (1994). Distinctions between two phonological output deficits. *Applied Psycholinguistics, 15*, 75–95.

Levy, B. A. (1971). The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behaviour, 10*, 123–132.

Martin, R. C., & Breedin, S. D. (1992). Dissociations between speech perception and phonological short term memory deficits. *Cognitive Neuropsychology, 9*, 509–534.

Martin, N., Dell, G. S., Saffran, E. M., & Schwarz, M. F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language, 47*, 609–660.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33*, 83–111.

McCarthy, R., & Warrington, E. K. (1984). A two-route model of speech production. Evidence from aphasia. *Brain, 107*, 483–485.

McKenna, P., & Warrington, E. K. (1983). *Graded naming test*. Windsor, UK: NFER-Nelson.

Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In D. A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Common processes in listening, speaking, reading and writing*. London: Academic Press.

Nickels, L., & Cole-Virtue, J. (2004). Reading tasks from PALPA: How do controls perform on visual lexical decision, homophony, rhyme, and synonym judgements? *Aphasiology, 18*(2), 103–126.

Nickels, L., Howard, D., & Best, W. (1997). Fractionating the articulatory loop: Dissociations and associations in phonological recoding in aphasia. *Brain and Language, 56*, 161–182.

Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition, 92*, 1–12.

Riddoch, M. J., & Humphreys, G. W. (1993). *BORB: Birmingham object recognition battery*. Hove,UK: Psychology Press.

Romani, C. (1992). Are there distinct input and output buffers? Evidence from a brain-damaged patient. *Language and Cognitive Processes, 7*, 131–162.

Shallice, T., & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia, 15*, 729–735.

Shallice, T., Rumiati, R. I., & Zadini, A. (2000). The selective impairment of the phonological output buffer. *Cognitive Neuropsychology, 17*, 417–546.

Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar, & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 11–53). Cambridge, UK: Cambridge University Press.

Tree, J. J., & Kay, J. (2015). Longitudinal assessment of short-term memory deterioration in a logopenic variant primary progressive aphasia with post-mortem confirmed Alzheimer's Disease pathology. *Journal of Neuropsychology, 9*(2), 184–202.

Tree, J. J., Longmore, C., & Besner, D. (2011). Orthography, phonology, short-term memory and the effects of concurrent articulation on rhyme and homophony judgements. *Acta Psychologica, 136*(1), 11–19.

Tree, J. J., Longmore, C., Majerus, S., & Evans, N. (2011). SKOOL versus ZOOL: Effects of orthographic and phonological long-term memory on nonword immediate serial recall. *Memory, 19*(5), 487–500.

Vallar & Papagno. (2002). Neuropsychological impairments of short-term memory. In Baddeley, Kopelman & Wilson (Ed.), *The handbook of memory disorders*. Sussex, UK.

Vallar, G., & Baddeley, A. D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology, 1*, 121–142.

Vallar, G., Di Betta, A. M., & Silveri, M. (1997). The phonological short-term store rehearsal system: Patterns of impairment and neural correlates. *Neuropsychologia, 35*, 795–812.

Warrington, E. K. (1997). The graded naming test: A restandardisation. *Neuropsychological Rehabilitation, 7*, 143–146.

Warrington, E. K., & James, M. (1991). *The visual object and Space perception battery*. Bury St Edmunds, UK: Thames Valley Test Company.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory-verbal short-term memory. *Brain, 92*, 885–896.

Warrington, E. K., & Shallice, T. (1972). Neuropsychological evidence of visual storage in short term memory tasks. *Quarterly Journal of Experimental Psychology, 24*, 30–40.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology, 50*(2), 439–456.

Please cite this article in press as: Tree, J. J., & Playfoot, D., How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM), Cortex (2018), https://doi.org/10.1016/j.cortex.2018.08.008