What is the relationship between phonological short-term memory and speech processing?

Charlotte Jacquemot¹ and Sophie K. Scott²

¹Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS-ENS-CNRS, 46 rue d'Ulm, 75005 Paris, France

²Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK

Corresponding author: Jacquemot, C. (charlotte.jacquemot@ens.fr).

Traditionally, models of speech comprehension and production do not depend on concepts and processes from the phonological short-term memory (pSTM) literature. Likewise, in working memory research, pSTM is considered to be a language-independent system that facilitates language acquisition rather than speech processing *per se*. We discuss couplings between pSTM, speech perception and speech production, and we propose that pSTM arises from the cycling of information between two phonological buffers, one involved in speech perception and one in speech production. We discuss the specific role of these processes in speech processing, and argue that models of speech perception and production, and our understanding of their neural bases, will benefit from incorporating them.

Introduction

The model of phonological short-term memory (pSTM) proposed by Baddeley *et al.* in 1984 [1] includes two components: a phonological buffer or store that can hold memory traces for a few seconds, and a subvocal rehearsal process used to refresh memory traces (Figure 1a). In this paper we address the relationship between pSTM and speech processing; although pSTM performance is influenced by phonological, lexical and semantic factors (Box 1), speech processing and pSTM have generally been studied separately, and are considered to be independent systems that can be selectively damaged [2]. Furthermore, models of speech processing do not generally identify a central role for pSTM [3,4]. Research has failed to reveal any direct correlation between speech comprehension and pSTM [5], and it has been proposed that pSTM has evolved as a language learning device [6].

We present data that support the argument that pSTM should be integrated in models of speech processing. In our approach, pSTM is an emergent property of the cycling of information between two phonological buffers involved in speech perception and production (Figure 1b). We first demonstrate commonalities between the mechanisms involved in speech perception and production and pSTM; we then discuss the role of phonological buffers; and finally we outline evidence for a direct role of each component of pSTM in speech processing.

How does speech affect pSTM performance?

Speech perception and the phonological buffer

The phonological buffer is assumed to store verbal information transiently, and verbal information seems to

be stored independently of non-verbal information. Recall performance is better with speech sounds than non-speech sounds [7] and recall of speech and non-speech sounds can be selectively damaged in patients [8]. When suffixes are added to the ends of speech and non-speech sequences, the suffix has a more deleterious effect on performance when it is of the same nature (speech or non-speech) as the sounds to be recalled [9].

The phonological nature of speech stimuli also affects pSTM performance. Both controls and brain-damaged patients show the 'phonological similarity effect': stimuli that are phonologically dissimilar are recalled better than similar ones [1]. Furthermore, subjects have difficulties in recalling sequences of stimuli that differ only for non-native language contrasts. For instance, the recall performance of French subjects worsens with sequences of stimuli that differ in stress location whereas Spanish subjects perform well (unlike Spanish, stress location in French is predictable and French speakers need not code stress in their phonological representations) [10]. These results suggest that native language properties influence recall performance and that the code used to store the stimuli is phonological in nature [11]. As the buffer capacity depends on these phonological properties, pSTM appears to be intimately connected to the speech perception system.

Speech production and the subvocal rehearsal component

pSTM also interacts closely with the speech production system. For example, pSTM performance is influenced by the length of stimuli: sequences of short words are recalled better than sequences of long words. This word length effect (WLE) was attributed to the longer rehearsal times for longer words [1] but the phonological complexity of the articulatory plan associated with the stimuli also affects recall performance [12]. The WLE is observed when spoken output is not required, meaning that it results not only from delay during output [13]. The WLE is abolished under articulatory suppression, suggesting that this effect is linked to the subvocal rehearsal component of pSTM [1]. pSTM performance also depends on constraints imposed by the subvocal production system and by how quickly speech stimuli can be produced subvocally (i.e. internally produced, without any spoken output) [14]. Digit spans are larger in languages whose digits are fast to pronounce [15], an effect also observed in bilingual subjects and therefore not explained by individual or cultural differences [16]. Finally, children who

1 misarticulate particular phonemes (e.g. /w/ and /r/) without making perception errors still make errors in pSTM tasks reflecting their specific phoneme substitutions (such as *ring* for *wing*) even without spoken responses [17]. Taken together, these results suggest that pSTM is closely connected to the speech production system, even when no spoken output is required.

8 Thus, we have shown that pSTM involves processes overlapping with both speech perception and production. The phonological buffer is involved in the storage of the phonological input and subvocal rehearsal requires the inner production of a speech output. Our claim predicts that the effects typically described as a signature of pSTM (phonological similarity effect, length effect, articulatory suppression effect) should be similarly observed in all languages, including sign language (Box 2).

17 One or two phonological buffers?

18 Debate continues as to whether there are two separate phonological buffers in perception and in production, or only one perceptual buffer [18]. Whereas the phonological input buffer was postulated in the first model of pSTM [1], a potential phonological output buffer appeared later in some models [2,19] but its role in the speech production system remains unspecified. However, a recent study of experimentally induced 'slips of the tongue' found a correlation between speech production errors and STM performance, suggesting a phonological output buffer involved in both STM and speech production [20].

29 Data from normal and brain-damaged subjects point to separate phonological input and output buffers. Investigations of the irrelevant sound effect – in which recall is impaired by concurrent or subsequent presentation of irrelevant acoustic material – have shown that the two phonological buffers are affected differently. The output buffer is disrupted more when irrelevant speech stimuli are similar to the material to be remembered, which is not the case for the input buffer [20,21]. In neuropsychological studies the distinction between input and output buffers relies on dissociations between STM tasks that do not require spoken output, such as list matching tasks (i.e. the subject has to compare two sequences of items) and those requiring speech production (i.e. the subject has to produce the memorized stimuli). Allport [22] compared the performance of two patients with severe restrictions of span in immediate serial list recall. In the matching span task, one patient performed as poorly as in spoken list recall, whereas the other patient showed good performance with lists much longer than those he could repeat. Allport suggested that the first patient had an impaired input buffer whereas the second had an impaired output buffer [23–26].

52 In the speech-processing field, there has also been interest in whether phonological representations are shared by perception and production systems or whether they are distinct. Neuropsychological evidence suggests that phonological representations in perception and production are distinct and connected by two processes, one that converts phonological input into phonological output and one that converts phonological output into phonological input [27]. Repetition of pseudowords uses

61 the conversion of phonological input into phonological output. Conversely, imagining the sound of a word or pseudoword subvocally involves the transformation of a phonological output into a phonological input without speech production. Thus phonemes can be identified in a picture name, without overtly naming the picture. These two conversion mechanisms can be selectively damaged in aphasic patients [27]. In addition, input and output buffers can be impaired in the absence of an explicit phonological encoding or decoding deficit [26,28,29].

71 Overall, these findings support the existence of two separate buffers that store phonological input and output transiently. We propose that pSTM arises from the recruitment of these two buffers and the cycling of information between them (Figure 1b). In our model, pSTM performance depends on the performance of both buffers and on the ability to convert information between them. Thus, a recall task with repetition involves both phonological buffers and their connections, whereas a matching task could involve only the phonological input buffer. With a damaged phonological input buffer, patients are impaired in all recall tasks (repetition and matching tasks) involving auditory pseudowords [1,23,26] whereas patients with deficits in the phonological output buffer perform relatively well on the matching task, but poorly on list repetition [24,30]. This model also predicts that a deficit of one of the conversion processes (from phonological input into output or vice versa) could occur, independent of any deficit of the input and output buffers, and would lead to the inability to use the phonological output buffer in a pSTM task even if the latter is not damaged [25].

93 What is the role of pSTM in speech processing?

94 Models of speech processing do not emphasize the role of pSTM, although both perception and production systems arguably require some buffer capacities to be operative. When hearing a sentence, there are no clear and reliable cues to the location of word boundaries and the system needs to segment a continuous speech stream (phonological input) into words. Mechanisms that could perform segmentation have been proposed and implemented in connectionist models, such as the competition or selection processes [3]. These mechanisms involve maintaining phonological representations that could encompass several syllables and that therefore require transient storage.

107 When producing a sentence, the elaboration of phonological output depends on the context in which words will appear – each word being influenced by the previous and following ones, and the syllabification boundaries are determined by various syntactic and morphological conditions [4]. Thus, to represent the syllables that compose the words, some buffer capacities are arguably necessary, although they are infrequently described or specified.

116 In the memory literature, patients with pSTM deficits and problems in comprehending semantically or syntactically complex sentences were reported in the early 1980s [5] with the claim that pSTM was necessary for semantic and syntactic processing. However, recent

1 studies suggest that the link between pSTM and syntactic
2 and semantic processing is more complex; some patients
3 with severely reduced verbal span showed preserved
4 sentence comprehension skills or normal abilities to
5 process syntactically complex sentences, whereas patients
6 with substantial speech perception deficits might well
7 have relatively well-preserved pSTM [5]. This work
8 suggested that pSTM has no direct role in speech
9 perception [6]. In the next sections we outline evidence for
10 processes that could nonetheless be common to both
11 speech processing and pSTM.

12 *The phonological input buffer*

13 There is now evidence that the phonological input buffer is
14 important in speech perception – specifically, a recent
15 study investigated its role in *sentence* perception [29].
16 Unlike previous studies that addressed the semantic and
17 syntactic processing of sentences [5], this study focused on
18 the phonological processing of sentences in two patients
19 who showed specific phonological input buffer deficits (e.g.
20 poor at recall tasks that involve pseudowords). Both
21 patients were impaired in comprehension when they were
22 required to process phonological details *within* a sentence.
23 Thus, although the patients could discriminate isolated
24 words (e.g. *coupure* and *couture*) that differ by only one
25 phoneme, they confused these same words within a
26 sentence context. In sentences, multiple syllables have to
27 be stored to compute lexical segmentation and resolve
28 phonological ambiguities [31]; the phonological input
29 buffer might be involved in this temporary storage [29],
30 and damage to this buffer seems to impair the
31 performance of these patients in the phonological analysis
32 of sentences.

33 *The phonological output buffer*

34 In speech-processing models, phonological output
35 comprises ordered phonological elements, strung together
36 to form syllables. The syllabification of a word within a
37 sentence is contextually influenced both segmentally (by
38 co-articulation and phonological rules) and prosodically
39 (by syntax) [32]. Syllabification runs over word boundaries
40 and uses the phonological forms of both previous and
41 upcoming words. We argue that the phonological output
42 buffer might perform this storage. Neuropsychological
43 patients with phonological output buffer deficits also make
44 speech production errors, such as substitution, insertion,
45 deletion and transposition of phonemes in all tasks
46 requiring speech output [28,33].

47 *The conversion of phonological input into phonological output*

48 Connections between phonological input and output allow
49 the repetition of auditory stimuli without lexical or
50 semantic-conceptual processing – for example, the
51 repetition of pseudowords or words from a foreign
52 language. In young children, significant correlations have
53 been found between performance in pseudoword repetition
54 and their vocabulary in both native and foreign languages
55 [34]. These data suggest that the transformation of
56 phonological input into output plays a role in learning new
57 words. Neuropsychological data confirm this: patients
58 with pseudoword repetition deficits are impaired in
59 learning foreign words [35]. Furthermore, data from

60 normal controls show that the acquisition of foreign
61 language vocabulary is disrupted by articulatory
62 suppression (unless the material allows semantic
63 associations to be created) [36].

64 *The conversion of phonological output into phonological input*

65 People can detect and correct errors of internal speech
66 production before producing them aloud [37]. It has been
67 proposed that monitoring errors in ‘internal speech’
68 involves the speech perception system, through an
69 internal loop between production and perception systems
70 [38], and experimental evidence suggests that the
71 monitoring of internal production operates on a syllabified
72 phonological output, more abstract than a phonetic
73 representation [39]. In our model (Figure 1b), this
74 corresponds to phonological output. We suggest that the
75 conversion of phonological output into phonological input
76 forms the internal loop of the monitoring system. We
77 predict that patients with impaired conversion processes
78 should spontaneously produce more speech errors than
79 controls: a prediction that needs to be tested. Notably,
80 patients with such deficits are impaired in tasks that
81 require them to internally ‘hear’ inner production, such as
82 visual rhyme judgment [25,26], consistent with our model.
83 Overall, these data suggest that the components
84 described in our model each have specific roles in speech
85 processing. We propose that pSTM is an emergent
86 property of the recruitment of these components. This
87 implies that the neural correlates of speech processing and
88 pSTM overlap (Box 3).

89 **Conclusion**

90 Buffering is essential when interacting processes function
91 on different timescales, and both speech production and
92 perception require that activation be maintained at
93 various points in processing. We argue that the
94 recruitment of buffers involved at the phonological level in
95 perception and production and the process of cycling of
96 information between the two buffers constitute what is
97 called pSTM. We have specified how pSTM could be
98 integrated within a model of speech processing and
99 identified the role of each component recruited for pSTM
100 in speech processing. Finally, the two conversion processes
101 provide a mechanism for the development and
102 stabilization of connections between the phonological
103 input and output and could be involved in the
104 maintenance of the ‘parity’ between the phonological
105 representations in input and output [40] (Box 4).

106 **Acknowledgements**

107 This work was supported by ESRC (C.J.) and the Wellcome Trust (S.K.S.).

108 **References**

- 1 Baddeley, A. *et al.* (1984) Exploring the articulatory loop. *Q. J. Exp. Psychol. A* 36, 233–252
- 2 Baddeley, A. (2003) Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839
- 3 Gaskell, M.G. and Marslen-Wilson, W.D. (2002) Representation and competition in the perception of spoken words. *Cognit. Psychol.* 45, 220–266
- 4 Levelt, W.J. (1999) Models of word production. *Trends Cogn. Sci.* 3, 223–232
- 5 Martin, R. (2006) The neuropsychology of sentence processing: where do we stand? *Cogn. Neuropsychol.* 23, 74–95
- 6 Baddeley, A. *et al.* (1998) The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173

- 7 Semal, C. *et al.* (1996) Speech versus nonspeech in pitch memory. *J. Acoust. Soc. Am.* 100, 1132–1140
- 8 Berndt, R.S. and Mitchum, C.S. (1990) Auditory and lexical information sources in immediate recall: evidence from a patient with deficit to the phonological short-term store. In *Neuropsychological Impairments of Short-term Memory* (Vallar, G. and Shallice, T., eds), pp. 115–144, Cambridge University Press
- 9 Rowe, E.J. and Rowe, W.G. (1976) Stimulus suffix effects with speech and nonspeech sounds. *Mem. Cognit.* 4, 128–131
- 10 Dupoux, E. *et al.* (2001) A robust method to study stress 'deafness'. *J. Acoust. Soc. Am.* 110, 1606–1618
- 11 Thorn, A.S. and Gathercole, S.E. (2001) Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychon. Bull. Rev.* 8, 357–364
- 12 Caplan, D. *et al.* (1992) Articulatory and phonological determinants of word length effects in span tasks. *Q. J. Exp. Psychol. A* 45, 177–192
- 13 Baddeley, A. *et al.* (2002) Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Q. J. Exp. Psychol. A* 55, 353–369
- 14 Wilson, M. (2001) The case for sensorimotor coding in working memory. *Psychon. Bull. Rev.* 8, 44–57
- 15 Stigler, J.W. *et al.* (1986) Digit memory in Chinese and English: evidence for a temporally limited store. *Cognition* 23, 1–20
- 16 Murray, A. and Jones, D.M. (2002) Articulatory complexity at item boundaries in serial recall: the case of Welsh and English digit span. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 594–598
- 17 Locke, J.L. and Kutz, K.J. (1975) Memory for speech and speech for memory. *J. Speech Hear. Res.* 18, 176–191
- 18 Monsell, S. (1987) On the relation between lexical input and output pathways of speech. In *Language Perception and Production: Relationships Between Listening, Speaking, Reading and Writing* (Allport, A. *et al.*, eds), pp. 273–311, Academic Press
- 19 Vallar, G. (2006) Memory systems: the case of phonological short-term memory. A festschrift for cognitive neuropsychology. *Cogn. Neuropsychol.* 23, 135–155
- 20 Saito, S. and Baddeley, A. (2004) Irrelevant sound disrupts speech production: exploring the relationship between short-term memory and experimentally induced slips of the tongue. *Q. J. Exp. Psychol. A* 57, 1309–1340
- 21 Jones, D.M. and Macken, W.J. (1995) Phonological similarity in the irrelevant sound effect: within- or between-stream similarity. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 103–115
- 22 Allport, A. (1984) Auditory verbal short-term memory and conduction aphasia. In *Attention and Performance X: Control and Language Processes* (Bouma, H. and Bouwhuis, D.G., eds), pp. 351–364, Erlbaum
- 23 Martin, R. *et al.* (1994) Language processing and working memory: neuropsychological evidence for separate phonological and semantic capacities. *J. Mem. Lang.* 33, 83–111
- 24 Romani, C. (1992) Are there distinct input and output buffers? Evidence from an aphasic patient with an impaired output buffer. *Lang. Cogn. Process.* 7, 131–162
- 25 Nickels, L. *et al.* (1997) Fractionating the articulatory loop: dissociations and associations in phonological recoding in aphasia. *Brain Lang.* 56, 161–182
- 26 Howard, D. and Nickels, L. (2005) Separating input and output phonology: semantic, phonological, and orthographic effects in short-term memory impairment. *Cogn. Neuropsychol.* 22, 42–77
- 27 Jacquemot, C. *et al.* Breaking the mirror: asymmetrical disconnection between the phonological input and output codes. *Cogn. Neuropsychol.* DOI: 10.1080/02643290600683342
- 28 Cohen, L. and Bachoud-Lévi, A.C. (1995) The role of the output phonological buffer in the control of speech timing: a single case study. *Cortex* 31, 469–486
- 29 Jacquemot, C. *et al.* (2006) Misperception in sentences but not in words: speech perception and the phonological buffer. *Cogn. Neuropsychol.* 23, 949–971
- 30 Martin, R. *et al.* (1999) Independence of input and output phonology in word processing and short-term memory. *J. Mem. Lang.* 41, 3–29
- 31 Gaskell, M.G. and Marslen-Wilson, W. (2001) Lexical ambiguity and spoken word recognition: bridging the gap. *J. Mem. Lang.* 44, 325–349
- 32 Levelt, W.J. (2001) Spoken word production: a theory of lexical access. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13464–13471
- 33 Shallice, T. *et al.* (2000) The selective impairment of the phonological output buffer. *Cogn. Neuropsychol.* 17, 517–546
- 34 Gathercole, S.E. (1999) Cognitive approaches to the development of short-term memory. *Trends Cogn. Sci.* 3, 410–419
- 35 Baddeley, A. (1993) Short-term phonological memory and long-term learning: a single case study. *Eur. J. Cogn. Psychol.* 5, 129–148
- 36 Papagno, C. *et al.* (1991) Phonological short-term memory and foreign language vocabulary learning. *J. Mem. Lang.* 30, 331–347
- 37 Levelt, W.J.M. *et al.* (1999) A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–37
- 38 Morgan, J.L. and Wheeldon, L.R. (2003) Syllable monitoring in internally and externally generated English words. *J. Psycholinguist. Res.* 32, 269–296
- 39 Wheeldon, L. and Levelt, W.J. (1995) Monitoring the time course of phonological encoding. *J. Mem. Lang.* 34, 311–334
- 40 Hickok, G. and Poeppel, D. (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99
- 41 Cipolotti, L. *et al.* (1991) A specific deficit for numbers in a case of dense acalculia. *Brain* 114, 2619–2637
- 42 Jefferies, E. *et al.* (2004) A category-specific advantage for numbers in verbal short-term memory: evidence from semantic dementia. *Neuropsychologia* 42, 639–660
- 43 Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114
- 44 Thorn, A.S. *et al.* (2005) Redintegration and the benefits of long-term knowledge in verbal short-term memory: an evaluation of Schweickert's (1993) multinomial processing tree model. *Cognit. Psychol.* 50, 133–158
- 45 Freedman, M.L. and Martin, R.C. (2001) Dissociable components of short-term memory and their relation to long-term learning. *Cogn. Neuropsychol.* 18, 193–226
- 46 Hulme, C. *et al.* (1997) Word-frequency effects on short-term memory tasks: evidence for a redintegration process in immediate serial recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 1217–1232
- 47 Gathercole, S.E. *et al.* (2001) Dissociable lexical and phonological influences on serial recognition and serial recall. *Q. J. Exp. Psychol. A* 54, 1–30
- 48 Corina, D.P. and Sandler, W. (1993) On the nature of phonological structure in sign language. *Phonology* 10, 165–207
- 49 Wilson, M. and Emmorey, K. (1997) A visuospatial 'phonological loop' in working memory: evidence from American Sign Language. *Mem. Cognit.* 25, 313–320
- 50 Baddeley, A. (1986) *Working Memory*, Oxford University Press
- 51 Wilson, M. and Emmorey, K. (1998) A 'word length effect' for sign language: further evidence for the role of language in structuring working memory. *Mem. Cognit.* 26, 584–590
- 52 Wager, T.D. and Smith, E.E. (2003) Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274
- 53 Wise, R.J. *et al.* (2001) Separate neural subsystems within 'Wernicke's area'. *Brain* 124, 83–95
- 54 Hickok, G. *et al.* (2003) Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* 15, 673–682
- 55 Jacquemot, C. *et al.* (2003) Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J. Neurosci.* 23, 9541–9546

Buchsbaum, B. *et al.* (2005) Neural substrates for verbal working memory in deaf signers: fMRI study and lesion case report. *Brain Lang.* 95, 265–272

Chein, J.M. and Fiez, J.A. (2001) Dissociation of verbal working memory system components using a delayed serial recall task. *Cereb. Cortex* 11, 1003–1014

Wise, R.J. *et al.* (1999) Brain regions involved in articulation. *Lancet* 353, 1057–1061

Catani, M. *et al.* (2005) Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16

Matsumoto, R. *et al.* (2004) Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain* 127, 2316–2330

TRENDS in Cognitive Sciences

16
17

Figure 1. Models of pSTM. In (a) and (b), pSTM appears in red. (a) A model of pSTM, as proposed by Baddeley *et al.* [1]. (b) Our proposed model of pSTM, integrated in a model of speech processing (only feedforward connections are reported). In this model, speech comprehension involves three steps. First, phonological decoding is defined as the translation of acoustic information into discrete segmental categories that belong to the language, i.e. the phonological input. Second, lexical recognition, which itself is formed of two components: word form selection and lexical retrieval. Word form selection involves the comparison of the speech sounds of the phonological input with those stored in lexical entries and in selecting the best matched word. Lexical retrieval results from the recovery of semantic information attached to the selected lexical entry. Speech production involves the same steps but in reverse order; lexical selection consists of selecting the word that corresponds to the information we want to express, and word form retrieval corresponds to the recovery of the phonological form associated with the word, i.e. the phonological output. Finally, phonological encoding involves the activation of the motor programme for producing the word. In our model, phonological representations in perception and production are distinct. At the lexical word form level, there might be two distinct representations for comprehension and production or a common one; this issue is outside the scope of this article. In this model, pSTM is composed of the two buffers dedicated to phonological processing in perception and production, and the mechanisms that convert phonological input information into output and vice versa. The input and output buffers are devices allowing the temporary maintenance of phonological representations in input and output.

29 Box 1. Digit span and pSTM performance

Traditionally, pSTM performance is evaluated by assessing digit span, which raises some problems. Digits seem to have a specific status compared to other semantic categories [41]. Studies have reported patients with a semantic dementia that have relatively preserved STM capacities when tested on digit task recall compared to other linguistic materials [42], suggesting that digits are processed differently from other semantic categories in STM.

A more general problem is raised by the use of lexical items in pSTM tasks. Using words – which by definition have lexico-semantic and conceptual representations – involves the recruitment of long-term memory (LTM), which biases the pSTM performance [43]. Studies on normal controls have demonstrated that memory span is influenced by phonological, lexical and semantic factors [44]. These observations have led to multibuffer models of verbal STM; evidence from neuropsychological studies supports the conclusion that not only is semantic and phonological information stored in STM but also the capacities for retaining the two types of information are separable. Indeed, studies have demonstrated that dissociations might be obtained between patients' ability to retain phonological and semantic information, suggesting the presence of a buffer specific to the lexical level in addition to the buffer capacities observed at the phonological level [23,45].

Another explanation for the effects of phonological, lexico-semantic, conceptual knowledge on pSTM is the redintegration process [46]. The term redintegration is used to describe the process by which, before output, incomplete phonological traces held in STM are reconstructed (redintegrated) by using knowledge relating to the phonological, lexico-semantic and conceptual properties of specific items. Reconstructive processes can occur either during storage (for instance, the effect of phonological properties during phonological decoding) or at rehearsal or retrieval (the effect of lexico-semantic and conceptual properties) [46,47].

Although there is little agreement about the source of the effect, long-term representations clearly influence recall performance in span tasks. Therefore, span tasks using pseudowords provide a more sensitive measure of *phonological* STM capacity than tasks involving real words because of the absence of information associated with pseudowords.

Box 2. Sign language and pSTM

Sign languages are highly complex and organize elementary, meaningless units into meaningful semantic units. The properties of the meaningless units are similar to the phonology of spoken languages (e.g. hierarchically organized feature classes, autosegmental representations, phonological rules, phonological hierarchy) [48]. For this reason, linguists have broadened the term 'phonology' to refer to the 'patterning of the formational units' of any natural language.

Signs are classically analysed into four basic phonological components: handshape, location in space, movement, and palm orientation (similar to features such as voicing, manner and place of articulation in spoken language). Signs that share at least one of these components are called similar signs. Signs that involve large movements are called long signs compared to those that involve no change of hand location (short signs).

Evidence suggests that in sign languages, as in spoken languages, information is stored in a phonological code. Tested with sequences of signed stimuli, deaf signers show a similarity effect: recall performance is lower for sequences composed of similar signs than for those composed of dissimilar signs. Tested on similar tasks but with nameable pictures instead of signed stimuli, deaf signers show a similarity effect suggesting that pictures are recoded into a phonological form. But this similarity effect disappears under manual 'articulatory' suppression (involving simple repetitive movements of the hands) [49]. Thus, in spoken and signed languages, under indirect presentation (pictures or written names), suppression abolishes the phonological similarity effect [49,50]. It seems therefore that a recoding process is needed to translate picture material into a phonological code and this is not the case when the stimuli to be remembered are signed or spoken stimuli.

As in spoken language, deaf signers' performance is lower when signed stimuli to be remembered are long rather than short [51]. This 'sign length' effect is similar to the 'word length' effect found in hearing subjects: under manual articulatory suppression the length effect is abolished suggesting that the length effect originates from the rehearsal process.

Taken together, these results suggest that signed pSTM consists of a buffer that stores information using the phonological structure of the language, and a submanual rehearsal process that seems to operate like the subvocal rehearsal process described in hearing subjects [14]. These data suggest that, whatever the structure of *phonological* STM (spoken or signed), the same processes are involved, and reflect a common mechanism.

Box 3. Neural basis of pSTM

Several neuroimaging studies have investigated the neural basis of verbal working memory. The activated areas included the left posterior parietal cortex (BA 40), Broca's area (BA 44/45), the left premotor area (BA 6) and the left supplementary area (SMA, BA 6) [52]. The left BA 40 was identified as the locus of the storage component of the pSTM, and Broca's area as involved in the rehearsal component.

As we propose that pSTM involves processes devoted to speech perception and production, the neural correlates of pSTM should also be activated in speech perception and production tasks. We should also expect two distinct regions to be related to the phonological input and output buffers. In perception, the left temporo-parietal junction, including the posterior superior temporal sulcus, supramarginal gyrus and medial planum temporale (Figure I; pSTS, SMG and MPT), has been observed in tasks involving the temporary storage of verbal input in signed and spoken languages [53–56]. Activity across or between these regions could be associated with functions of the phonological input buffer.

With respect to the phonological output buffer, the activation of the left inferior frontal gyrus (Figure I; LIFG) encompassing the inferior precentral gyrus (BA 6) is observed in tasks involving the storage of a phonological output [57]. Moreover, the left anterior insula (Figure I; I), known to be involved in the control of speech output, has shown sensitivity to speech perception [58], suggesting that these areas beyond the left frontal inferior gyrus and including part of the middle frontal and inferior precentral gyri might correspond to the function of an output buffer.

Using tractography techniques, Catani *et al.* [59] have shown that these areas are connected to the superior temporal region through two pathways: a direct pathway and an indirect pathway via the inferior parietal region. Moreover, findings from a cortico-cortical-evoked potential study revealed a bidirectional connection between frontal and temporo-parietal language areas [60]. This suggests that, unlike the classical Wernicke–Geschwind model, the language areas involved in production and perception are reciprocally connected. These bidirectional connections between regions associated with the phonological output buffer and the phonological input buffer are a good candidate for the neural substrates of the mechanism that converts information between the two phonological buffers.

TRENDS in Cognitive Sciences

1
2

3 Figure 1. Cortical candidate neural regions involved in pSTM as speech input or output buffers. The regions in red correspond to a speech input buffer system in the posterior
4 superior temporal sulcus (pSTS), the supramarginal gyrus (SMG) and the medial planum temporal (MPT). The regions in green correspond to a speech output buffer system,
5 potentially seen in the left inferior frontal gyrus (LIFG) extending into the inferior motor cortex and in the anterior insula (I). These regions are shown schematically using coronal
6 brain slices on the left of the figure, and the cortical surface of the left hemisphere on the right.

7 **Box 4. Questions for future research**

- 8 • The ability to map a speech input to a speech output has arguably been a crucial step in language development. Does it follow therefore that
9 language and pSTM should have appeared and evolved conjointly? Is pSTM a prerequisite for the development of such a complex information
10 structure as human language?
- 11 • Are there further buffers associated with further levels of speech processing (semantic, syntactic)?
- 12 • What is the nature of the representations at phonological input and output levels?
- 13 • How can we best specify the conversion procedures between phonological input and output?
- 14