



HAL
open science

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

Charlotte Jacquemot, Sophie Scott

► To cite this version:

Charlotte Jacquemot, Sophie Scott. What is the relationship between phonological short-term memory and speech processing?. Trends in Cognitive Sciences, 2006, 10 (11), pp.480-486. 10.1016/j.tics.2006.09.002 . hal-02326818

HAL Id: hal-02326818

<https://hal.science/hal-02326818>

Submitted on 19 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 What is the relationship between 2 phonological short-term memory 3 and speech processing?

4 Charlotte Jacquemot¹ and Sophie K. Scott²

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

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

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

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

22 Introduction

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

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

49 How does speech affect pSTM performance?

50 *Speech perception and the phonological buffer*

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

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

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

79 *Speech production and the subvocal rehearsal component*

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

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

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

17 One or two phonological buffers?

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

29 Data from normal and brain-damaged subjects point to
30 separate phonological input and output buffers.
31 Investigations of the irrelevant sound effect – in which
32 recall is impaired by concurrent or subsequent
33 presentation of irrelevant acoustic material – have shown
34 that the two phonological buffers are affected differently.
35 The output buffer is disrupted more when irrelevant
36 speech stimuli are similar to the material to be
37 remembered, which is not the case for the input buffer
38 [20,21]. In neuropsychological studies the distinction
39 between input and output buffers relies on dissociations
40 between STM tasks that do not require spoken output,
41 such as list matching tasks (i.e. the subject has to compare
42 two sequences of items) and those requiring speech
43 production (i.e. the subject has to produce the memorized
44 stimuli). Allport [22] compared the performance of two
45 patients with severe restrictions of span in immediate
46 serial list recall. In the matching span task, one patient
47 performed as poorly as in spoken list recall, whereas the
48 other patient showed good performance with lists much
49 longer than those he could repeat. Allport suggested that
50 the first patient had an impaired input buffer whereas the
51 second had an impaired output buffer [23–26].

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

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

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

93 What is the role of pSTM in speech processing?

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

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

116 In the memory literature, patients with pSTM deficits
117 and problems in comprehending semantically or
118 syntactically complex sentences were reported in the early
119 1980s [5] with the claim that pSTM was necessary for
120 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.).

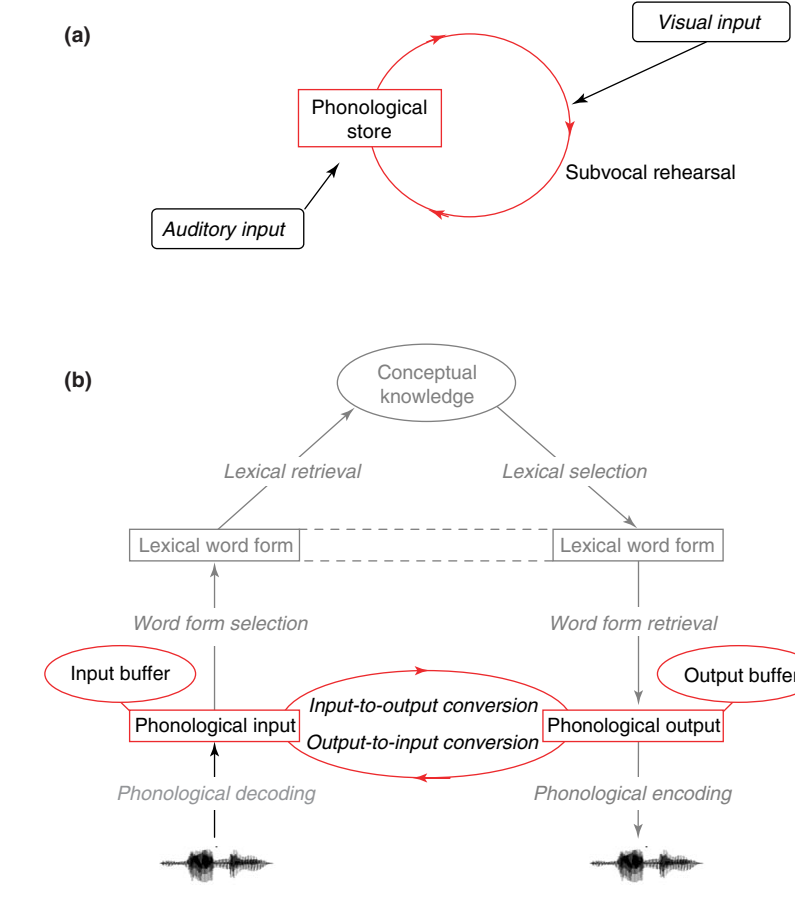
108 **References**

- 109 1 Baddeley, A. *et al.* (1984) Exploring the articulatory loop. *Q. J. Exp. Psychol. A* 36, 233–252
- 110 2 Baddeley, A. (2003) Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839
- 111 3 Gaskell, M.G. and Marslen-Wilson, W.D. (2002) Representation and competition in the perception of spoken words. *Cognit. Psychol.* 45, 220–266
- 112 4 Levelt, W.J. (1999) Models of word production. *Trends Cogn. Sci.* 3, 223–232
- 113 5 Martin, R. (2006) The neuropsychology of sentence processing: where do we stand? *Cogn. Neuropsychol.* 23, 74–95
- 114 6 Baddeley, A. *et al.* (1998) The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173

1	7 Semal, C. <i>et al.</i> (1996) Speech versus nonspeech in pitch	74	30	Martin, R. <i>et al.</i> (1999) Independence of input and	
2	memory. <i>J. Acoust. Soc. Am.</i> 100, 1132–1140	75	output phonology in word processing and short-term memory. <i>J.</i>		
3	8 Berndt, R.S. and Mitchum, C.S. (1990) Auditory and lexical	76	<i>Mem. Lang.</i> 41, 3–29		
4	information sources in immediate recall: evidence from a patient	77	31	Gaskell, M.G. and Marslen-Wilson, W. (2001) Lexical	
5	with deficit in the phonological short-term store. In	78	ambiguity and spoken word recognition: bridging the gap. <i>J. Mem.</i>		
6	<i>Neuropsychological Impairments of Short-term Memory</i> (Vallar, G.	79	<i>Lang.</i> 44, 325–349		
7	and Shallice, T., eds), pp. 115–144, Cambridge University Press	80	32	Levelt, W.J. (2001) Spoken word production: a theory	
8	9 Rowe, E.J. and Rowe, W.G. (1976) Stimulus suffix effects with	81	of lexical access. <i>Proc. Natl. Acad. Sci. U. S. A.</i> 98, 13464–13471		
9	speech and nonspeech sounds. <i>Mem. Cognit.</i> 4, 128–131	82	33	Shallice, T. <i>et al.</i> (2000) The selective impairment of	
10	10 Dupoux, E. <i>et al.</i> (2001) A robust method to study	83	the phonological output buffer. <i>Cogn. Neuropsychol.</i> 17, 517–546		
11	stress 'deafness'. <i>J. Acoust. Soc. Am.</i> 110, 1606–1618	84	34	Gathercole, S.E. (1999) Cognitive approaches to the	
12	11 Thorn, A.S. and Gathercole, S.E. (2001) Language	85	development of short-term memory. <i>Trends Cogn. Sci.</i> 3, 410–419		
13	differences in verbal short-term memory do not exclusively	86	35	Baddeley, A. (1993) Short-term phonological memory	
14	originate in the process of subvocal rehearsal. <i>Psychon. Bull. Rev.</i>	87	and long-term learning: a single case study. <i>Eur. J. Cogn. Psychol.</i>		
15	8, 357–364	88	5, 129–148		
16	12 Caplan, D. <i>et al.</i> (1992) Articulatory and phonological	89	36	Papagno, C. <i>et al.</i> (1991) Phonological short-term	
17	determinants of word length effects in span tasks. <i>Q. J. Exp.</i>	90	memory and foreign language vocabulary learning. <i>J. Mem. Lang.</i>		
18	<i>Psychol. A</i> 45, 177–192	91	30, 331–347		
19	13 Baddeley, A. <i>et al.</i> (2002) Is the word length effect in	92	37	Levelt, W.J.M. <i>et al.</i> (1999) A theory of lexical access	
20	STM entirely attributable to output delay? Evidence from serial	93	in speech production. <i>Behav. Brain Sci.</i> 22, 1–37		
21	recognition. <i>Q. J. Exp. Psychol. A</i> 55, 353–369	94	38	Morgan, J.L. and Wheeldon, L.R. (2003) Syllable	
22	14 Wilson, M. (2001) The case for sensorimotor coding	95	monitoring in internally and externally generated English words.		
23	in working memory. <i>Psychon. Bull. Rev.</i> 8, 44–57	96	<i>J. Psycholinguist. Res.</i> 32, 269–296		
24	15 Stigler, J.W. <i>et al.</i> (1986) Digit memory in Chinese	97	39	Wheeldon, L. and Levelt, W.J. (1995) Monitoring the	
25	and English: evidence for a temporally limited store. <i>Cognition</i> 23,	98	time course of phonological encoding. <i>J. Mem. Lang.</i> 34, 311–334		
26	1–20	99	40	Hickok, G. and Poeppel, D. (2004) Dorsal and ventral	
27	16 Murray, A. and Jones, D.M. (2002) Articulatory	100	streams: a framework for understanding aspects of the functional		
28	complexity at item boundaries in serial recall: the case of Welsh	101	anatomy of language. <i>Cognition</i> 92, 67–99		
29	and English digit span. <i>J. Exp. Psychol. Learn. Mem. Cogn.</i> 28,	102	41	Cipolotti, L. <i>et al.</i> (1991) A specific deficit for	
30	594–598	103	numbers in a case of dense acalculia. <i>Brain</i> 114, 2619–2637		
31	17 Locke, J.L. and Kutz, K.J. (1975) Memory for speech	104	42	Jefferies, E. <i>et al.</i> (2004) A category-specific	
32	and speech for memory. <i>J. Speech Hear. Res.</i> 18, 176–191	105	advantage for numbers in verbal short-term memory: evidence		
33	18 Monsell, S. (1987) On the relation between lexical	106	from semantic dementia. <i>Neuropsychologia</i> 42, 639–660		
34	input and output pathways of speech. In <i>Language Perception and</i>	107	43	Cowan, N. (2001) The magical number 4 in short-	
35	<i>Production: Relationships Between Listening, Speaking, Reading</i>	108	term memory: a reconsideration of mental storage capacity. <i>Behav.</i>		
36	<i>and Writing</i> (Allport, A. <i>et al.</i> , eds), pp. 273–311, Academic Press	109	<i>Brain Sci.</i> 24, 87–114		
37	19 Vallar, G. (2006) Memory systems: the case of	110	44	Thorn, A.S. <i>et al.</i> (2005) Redintegration and the	
38	phonological short-term memory. A festschrift for cognitive	111	benefits of long-term knowledge in verbal short-term memory: an		
39	neuropsychology. <i>Cogn. Neuropsychol.</i> 23, 135–155	112	evaluation of Schweickert's (1993) multinomial processing tree		
40	20 Saito, S. and Baddeley, A. (2004) Irrelevant sound	113	model. <i>Cognit. Psychol.</i> 50, 133–158		
41	disrupts speech production: exploring the relationship between	114	45	Freedman, M.L. and Martin, R.C. (2001) Dissociable	
42	short-term memory and experimentally induced slips of the	115	components of short-term memory and their relation to long-term		
43	tongue. <i>Q. J. Exp. Psychol. A</i> 57, 1309–1340	116	learning. <i>Cogn. Neuropsychol.</i> 18, 193–226		
44	21 Jones, D.M. and Macken, W.J. (1995) Phonological	117	46	Hulme, C. <i>et al.</i> (1997) Word-frequency effects on	
45	similarity in the irrelevant sound effect: within- or between-	118	short-term memory tasks: evidence for a redintegration process in		
46	stream similarity. <i>J. Exp. Psychol. Learn. Mem. Cogn.</i> 21, 103–115	119	immediate serial recall. <i>J. Exp. Psychol. Learn. Mem. Cogn.</i> 23,		
47	22 Allport, A. (1984) Auditory verbal short-term	120	1217–1232		
48	memory and conduction aphasia. In <i>Attention and Performance X:</i>	121	47	Gathercole, S.E. <i>et al.</i> (2001) Dissociable lexical and	
49	<i>Control and Language Processes</i> (Bouma, H. and Bouwhuis, D.G.,	122	phonological influences on serial recognition and serial recall. <i>Q.</i>		
50	eds), pp. 351–364, Erlbaum	123	<i>J. Exp. Psychol. A</i> 54, 1–30		
51	23 Martin, R. <i>et al.</i> (1994) Language processing and	124	48	Corina, D.P. and Sandler, W. (1993) On the nature of	
52	working memory: neuropsychological evidence for separate	125	phonological structure in sign language. <i>Phonology</i> 10, 165–207		
53	phonological and semantic capacities. <i>J. Mem. Lang.</i> 33, 83–111	126	49	Wilson, M. and Emmorey, K. (1997) A visuospatial	
54	24 Romani, C. (1992) Are there distinct input and	127	'phonological loop' in working memory: evidence from American		
55	output buffers? Evidence from an aphasic patient with an	128	Sign Language. <i>Mem. Cognit.</i> 25, 313–320		
56	impaired output buffer. <i>Lang. Cogn. Process.</i> 7, 131–162	129	50	Baddeley, A. (1986) <i>Working Memory</i> , Oxford	
57	25 Nickels, L. <i>et al.</i> (1997) Fractionating the	130	University Press		
58	articulatory loop: dissociations and associations in phonological	131	51	Wilson, M. and Emmorey, K. (1998) A 'word length	
59	recoding in aphasia. <i>Brain Lang.</i> 56, 161–182	132	effect' for sign language: further evidence for the role of language		
60	26 Howard, D. and Nickels, L. (2005) Separating input	133	in structuring working memory. <i>Mem. Cognit.</i> 26, 584–590		
61	and output phonology: semantic, phonological, and orthographic	134	52	Wager, T.D. and Smith, E.E. (2003) Neuroimaging	
62	effects in short-term memory impairment. <i>Cogn. Neuropsychol.</i> 22,	135	studies of working memory: a meta-analysis. <i>Cogn. Affect. Behav.</i>		
63	42–77	136	<i>Neurosci.</i> 3, 255–274		
64	27 Jacquemot, C. <i>et al.</i> Breaking the mirror:	137	53	Wise, R.J. <i>et al.</i> (2001) Separate neural subsystems	
65	asymmetrical disconnection between the phonological input and	138	within 'Wernicke's area'. <i>Brain</i> 124, 83–95		
66	output codes. <i>Cogn. Neuropsychol.</i> DOI:	139	54	Hickok, G. <i>et al.</i> (2003) Auditory-motor interaction	
67	10.1080/02643290600683342	28	revealed by fMRI: speech, music, and working memory in area		
68	Cohen, L. and	140	Spt. <i>J. Cogn. Neurosci.</i> 15, 673–682		
69	Bachoud-Lévi, A.C. (1995) The role of the output phonological	141	55	Jacquemot, C. <i>et al.</i> (2003) Phonological grammar	
70	buffer in the control of speech timing: a single case study. <i>Cortex</i>	142	shapes the auditory cortex: a functional magnetic resonance		
71	31, 469–486	143	imaging study. <i>J. Neurosci.</i> 23, 9541–9546		
72	29 Jacquemot, C. <i>et al.</i> (2006) Misperception in	144			
73	sentences but not in words: speech perception and the phonological				
	buffer. <i>Cogn. Neuropsychol.</i> 23, 949–971				

1 56 Buchsbaum, B. *et al.* (2005) Neural substrates for 9
 2 verbal working memory in deaf signers: fMRI study and lesion 10
 3 case report. *Brain Lang.* 95, 265–272 11
 4 57 Chein, J.M. and Fiez, J.A. (2001) Dissociation of 12
 5 verbal working memory system components using a delayed serial 13
 6 recall task. *Cereb. Cortex* 11, 1003–1014 14
 7 58 Wise, R.J. *et al.* (1999) Brain regions involved in
 8 articulation. *Lancet* 353, 1057–1061

59 Catani, M. *et al.* (2005) Perisylvian language
 networks of the human brain. *Ann. Neurol.* 57, 8–16
 60 Matsumoto, R. *et al.* (2004) Functional connectivity
 in the human language system: a cortico-cortical evoked potential
 study. *Brain* 127, 2316–2330



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.

18 **Box 1. Digit span and pSTM performance**
 19 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.
 20 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].

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

6 Although there is little agreement about the source of the effect, long-term representations clearly influence recall performance in span tasks.
7 Therefore, span tasks using pseudowords provide a more sensitive measure of *phonological* STM capacity than tasks involving real words because of the
8 absence of information associated with pseudowords.
9

10 **Box 2. Sign language and pSTM**

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

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

18 Evidence suggests that in sign languages, as in spoken languages, information is stored in a phonological code. Tested with sequences of signed
19 stimuli, deaf signers show a similarity effect: recall performance is lower for sequences composed of similar signs than for those composed of dissimilar
20 signs. Tested on similar tasks but with nameable pictures instead of signed stimuli, deaf signers show a similarity effect suggesting that pictures are
21 recoded into a phonological form. But this similarity effect disappears under manual 'articulatory' suppression (involving simple repetitive movements of
22 the hands) [49]. Thus, in spoken and signed languages, under indirect presentation (pictures or written names), suppression abolishes the phonological
23 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
24 when the stimuli to be remembered are signed or spoken stimuli.

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

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

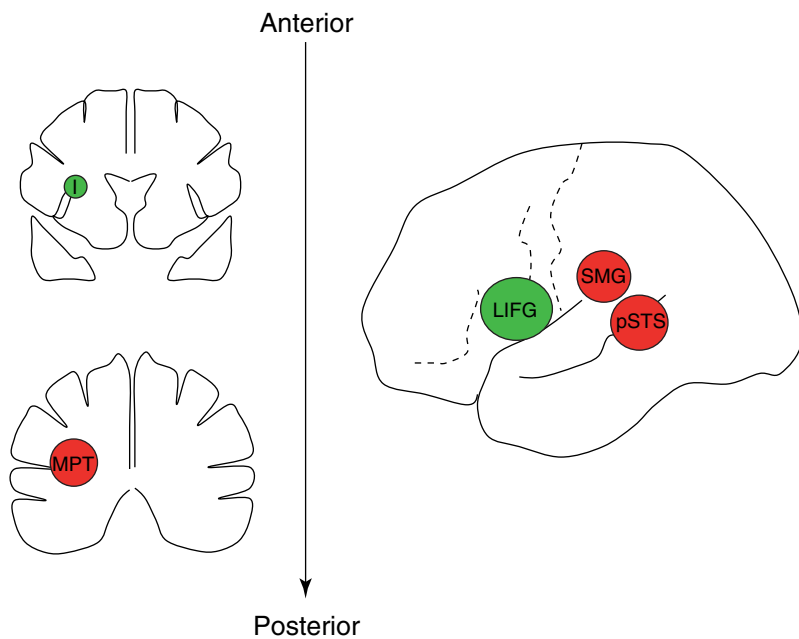
31 **Box 3. Neural basis of pSTM**

32 Several neuroimaging studies have investigated the neural basis of verbal working memory. The activated areas included the left posterior parietal cortex
33 (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
34 locus of the storage component of the pSTM, and Broca's area as involved in the rehearsal component.

35 As we propose that pSTM involves processes devoted to speech perception and production, the neural correlates of pSTM should also be activated in
36 speech perception and production tasks. We should also expect two distinct regions to be related to the phonological input and output buffers. In
37 perception, the left temporo-parietal junction, including the posterior superior temporal sulcus, supramarginal gyrus and medial planum temporale
38 (Figure 1; pSTS, SMG and MPPT), has been observed in tasks involving the temporary storage of verbal input in signed and spoken languages [53–56].
39 Activity across or between these regions could be associated with functions of the phonological input buffer.

40 With respect to the phonological output buffer, the activation of the left inferior frontal gyrus (Figure 1; LIFG) encompassing the inferior precentral gyrus
41 (BA 6) is observed in tasks involving the storage of a phonological output [57]. Moreover, the left anterior insula (Figure 1; I), known to be involved in the
42 control of speech output, has shown sensitivity to speech perception [58], suggesting that these areas beyond the left frontal inferior gyrus and including
43 part of the middle frontal and inferior precentral gyri might correspond to the function of an output buffer.

44 Using tractography techniques, Catani *et al.* [59] have shown that these areas are connected to the superior temporal region through two pathways: a
45 direct pathway and an indirect pathway via the inferior parietal region. Moreover, findings from a cortico-cortical-evoked potential study revealed a
46 bidirectional connection between frontal and temporo-parietal language areas [60]. This suggests that, unlike the classical Wernicke–Geschwind model,
47 the language areas involved in production and perception are reciprocally connected. These bidirectional connections between regions associated with
48 the phonological output buffer and the phonological input buffer are a good candidate for the neural substrates of the mechanism that converts
49 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