Short-term memory stages in sign vs. speech: The source of the serial span discrepancy

Hall, Matthew L.; Bavelier, Daphné

How to cite

HALL, Matthew L., BAVELIER, Daphné. Short-term memory stages in sign vs. speech: The source of the serial span discrepancy. In: Cognition, 2011, vol. 120, n° 1, p. 54–66. doi: 10.1016/j.cognition.2011.02.014

This publication URL: https://archive-ouverte.unige.ch/unige:102630

Publication DOI: 10.1016/j.cognition.2011.02.014

Cognition 120 (2011) 54–66

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Short-term memory stages in sign vs. speech: The source of the serial span discrepancy

Matthew L. Hall [a,*], Daphne Bavelier [b,*]

[a] Department of Psychology, UC San Diego, 9500 Gilman Dr., La Jolla, CA 92039-0109, United States
[b] Department of Brain and Cognitive Science, Meliora Hall -0268, U. of Rochester, Rochester NY 14627-0268, United States

ARTICLE INFO

Article history:
Received 6 January 2010
Revised 18 February 2011
Accepted 23 February 2011
Available online 29 March 2011

Keywords:
Short-term memory
Digit span
Sign language
Working memory
Phonological loop

ABSTRACT

Speakers generally outperform signers when asked to recall a list of unrelated verbal items. This phenomenon is well established, but its source has remained unclear. In this study, we evaluate the relative contribution of the three main processing stages of short-term memory – perception, encoding, and recall – in this effect. The present study factorially manipulates whether American Sign Language (ASL) or English is used for perception, memory encoding, and recall in hearing ASL-English bilinguals. Results indicate that using ASL during both perception and encoding contributes to the serial span discrepancy. Interestingly, performing recall in ASL slightly increased span, ruling out the view that signing is in general a poor choice for short-term memory. These results suggest that despite the general equivalence of sign and speech in other memory domains, speech-based representations are better suited for the specific task of perception and memory encoding of a series of unrelated verbal items in serial order through the phonological loop. This work suggests that interpretation of performance on serial recall tasks in English may not translate straightforwardly to serial tasks in sign language.

1. Introduction

The nature and efficacy of Deaf people's mental representations has been a question of enduring interest among researchers, clinicians, and educators alike. Decades of memory research have revealed overwhelming similarity between cognitive processes in deaf and hearing populations (Furth (1971) and Rudner, Andin, and Rönnberg (2009) for reviews). However, one task in which hearing subjects consistently outperform deaf subjects is the immediate serial recall of unrelated verbal items. This serial span discrepancy has been shown not only in American Sign Language (Bellugi, Klima, & Siple, 1975; Boutla, Supalla, Newport, & Bavelier, 2004; Hamilton & Holzman, 1989; Hanson, 1982; Hanson & Lichtenstein, 1990; Hoemann & Blama, 1992; Koo, Crain, LaSasso, & Eden, 2008; Krakow & Hanson, 1985; Lichtenstein, 1998; Pintner & Paterson, 1917; Wallace & Corballis, 1973), but also in Auslan (Logan, Maybery, & Fletcher, 1996), British Sign Language (Conrad, 1970; MacSweeney, Campbell, & Donlan, 1996), Italian Sign Language (Geraci, Gozzi, Papagno, & Cecchetto, 2008), Israeli Sign Language (Miller, 2007), and Swedish Sign Language (Rönnberg, Rudner, & Ingvar, 2004). Despite widespread agreement about the phenomenon itself, there is no consensus as to its source.

One possibility is that serial span in sign language is lower because sign language is visuospatial. It is well established that visuospatial span reaches a maximum of 4–5 in a variety of tasks (see Cowan, 2001), which is around the same span typically observed in signers. However, there is ample evidence that signers rely on a process that more closely resembles verbal coding than visuospatial coding. The strongest evidence comes from studies by Wilson and Emmorey (1997, 1998, 2003), who

* Corresponding authors.
E-mail addresses: matthall@ucsd.edu (M.L. Hall), daphne@bcs.rochester.edu (D. Bavelier).

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.02.014

demonstrated evidence in ASL signers of the four signature effects of verbal coding: phonological similarity, articulatory suppression, word length, and irrelevant sign. Such psychological evidence validates previous linguistic analyses demonstrating that signs do have sub-lexical structure (phonology) much like words do (Bellugi et al., 1975; Friedman, 1976; Klima & Bellugi, 1979). Thus, the notions of phonological coding and phonological similarity may be applied to sign as well as speech.

If indeed STM in sign relies on verbal coding rather than visuospatial coding, another possible source of the serial span discrepancy might be found in factors that are known to affect serial span tasks in speech. These include phonological similarity (Baddeley, Thomson, & Buchanan, 1975; Conrad & Hull, 1964), phonological complexity (Caplan, Rochon, & Waters, 1992), and the articulatory duration of the words used (Elliot, 1992; Ellis & Hennelly, 1980). Indeed, such factors were among the preferred explanations until Boutla and colleagues (2004) demonstrated a serial span discrepancy between ASL and English using items that were phonologically dissimilar, phonologically simple, and equally fast to articulate in ASL and English. Furthermore, they observed the same result in native ASL-English bilinguals, thus localizing the effect to the use of sign language, rather than to deafness. This work also ruled out an explanation for the serial span discrepancy in terms of reduced mnemonic ability in deaf participants. Although Wilson and Emmorey (2006) have challenged these findings on the grounds that Boutla et al. (2004) used letters in ASL but digits in English, the fact remains that a serial span discrepancy between sign and speech has been noted when comparing items matched for articulatory duration across languages whether they may be ASL digits vs. English digits (Bavelier, Newport, Hall, Supalla, & Boutla, 2008, Experiment 1), ASL fingerspelled letters vs. spoken English letters (Bavelier, Newport, et al., 2008, Experiment 3; Bavelier, Newport, Hall, Supalla, & Boutla 2006), or Italian Sign Language (LIS) nouns vs. spoken Italian nouns (Geraci et al., 2008).

The puzzle, then, persists. Given the evidence that signers use linguistic coding (rather than visuospatial), and having controlled for all known factors that influence linguistic STM span, at least in spoken language, where and when does the serial span discrepancy between signers and speakers originate? To address this issue we use here a within-subject design and systematically vary whether speech-based or sign-based representations are used as the various stages of a serial STM task unfold.

The first stage of short-term memory is perception, wherein a to-be-remembered stimulus is first encountered by the senses. It is well established that the auditory modality is better able to resolve temporal frequency than the visual modality, and signers do not differ from speakers in this respect (Poizner & Tallal, 1987). Some have suggested that these acoustic properties have trickle-down effects in later stages of cognitive processing and even higher cognition (Conway, Pisoni, & Kronenberger, 2009). Previous research of STM in speakers has found small but persistent advantages for spoken over written presentation (Penney, 1989), despite the fact that both engage speech-based coding. Thus, it seems plausible that the acoustics of speech could play a role in the serial span discrepancy. The critical manipulation at this stage concerns how stimuli are presented (e.g. sign, speech, print). Previous studies comparing serial recall of signed vs. spoken stimuli confounded presentation modality with hearing status (Bellugi et al., 1975; Hanson, 1982; Krakow & Hanson, 1985; Liben & Drury, 1977), or could not distinguish effects of presentation modality from those due to the internal code that the subjects used (Bavelier, Newport, et al., 2008; Boutla et al., 2004; Hamilton & Holzman, 1989; Hoemann & Blama, 1992; Koo et al., 2008; Rönnberg et al., 2004; Shand, 1982). The present study is the first to dissociate effects of perception from effects of relying on sign-based vs. speech-based memory encoding. To do so, we compared span size within-subjects and within-items as a function of whether the stimuli were presented in sign or in audio-visual speech, while controlling for the internal memory code used and language of recall. If the perceptual characteristics of the stimulus contribute to the serial span discrepancy, there should be an advantage for English presentation over ASL presentation.

The second stage of STM requires mapping the perceptual input onto a pre-existing mental representation. We refer to this stage as “encoding”, and the key question for this stage concerns the nature of the internal code that the subjects use to represent the stimulus in memory. For linguistic material, it is widely accepted that information is held in the phonological store and refreshed through an articulatory loop (Baddeley, 1986). Given the above evidence that it is possible for signers, like speakers, to rely on a sign-based phonological loop during verbal STM tasks, signed input should have ready access to an internal sign-based code. However, this equivalence in memory architecture does not necessarily entail equal mnemonic efficiency. Many researchers have suggested that a speech-based code might be better suited for serial recall than a sign-based code, and that this could be the source of the serial span discrepancy (Conrad, 1970; Hamilton & Holzman, 1989; Hanson, 1982; Koo et al., 2008; Krakow & Hanson, 1985; Lichtenstein, 1998; Miller, 2007). However, none of these studies tested this hypothesis by empirically manipulating the nature of the subjects’ internal codes. To do so in this study, we instructed participants to shadow the stimulus model by either signing or silently mouthing the to-be-remembered sequence before recalling from memory. We reasoned that overt (but silent) shadowing in speech or sign would be the surest way to engage speech-based or sign-based internal codes, respectively. The language used for shadowing was independent of language of presentation and of recall. If, as previous research suggests, the internal code that subjects use contributes to the serial span discrepancy, there should be an advantage for silent English shadowing over ASL shadowing.

The final stage of STM is recall, where the sequence is retrieved from memory and reproduced by the articulators. Forgetting can occur during the recall process (Dosher & Ma, 1998), especially when the to-be-recalled items take longer to produce. Because lexical signs tend to be longer to articulate than their spoken translations (Bellugi & Fischer, 1972; Klima & Bellugi, 1979), and to establish parity of output between deaf and hearing populations, most previous studies have used written recall. However, this raises several problems. For example, by forcing Deaf signers to recall items in written English, studies that use lexical stimuli (Bellugi et al., 1975; Hamilton & Holzman, 1989; Krakow & Hanson, 1985; Lichtenstein, 1998; Miller, 2007) require signers to make an extra mental translation that is not required of hearing English speakers, and this could be part of what reduces span. In addition, the manual motor acts involved in written output might interfere with a sign-based code more than they would interfere with a speech-based code. Finally, those studies which have allowed subjects to recall in sign (Boutla et al., 2004; Hanson, 1982; MacSweeney et al., 1996; Rönnberg et al., 2004) or in a modality of their choosing (Koo et al., 2008) cannot discriminate effects due to internal codes from effects that arise during recall. Thus, the impact of using sign vs. speech to perform recall is still unknown. In this study, we instructed participants to perform recall in either sign or speech. If using ASL to perform recall contributes to the serial span discrepancy, there should be an advantage for English recall over ASL recall.

The present study manipulates whether ASL or English is used for presentation, encoding (through shadowing), or recall. This $2 \times 2 \times 2$ within-subjects design calls for the testing of hearing ASL-English bilinguals. Because bilinguals vary in age of acquisition and proficiency between their two languages, we tested both early and late bilinguals of higher and lower proficiency, and included these as factors in our analysis.

2. Experiment 1

2.1. Method

2.1.1. Subjects

We tested 73 ASL-English bilinguals with normal hearing, and report data from 61 subjects, half of whom ($n = 31$) had Deaf parents and acquired both ASL and English as first languages from infancy (hereafter CODAs: Children of Deaf Adults). The other half ($n = 30$) were native English speakers who did not learn ASL until adolescence or later (Non-CODAs). Other demographics are given in Table 1.

To dissociate effects of proficiency from age of acquisition, we further divided the CODAs and Non-CODAs into higher-proficiency and lower-proficiency groups. In the absence of standardized measures of ASL fluency, we used professional interpreting experience as a proxy for proficiency. We reasoned that although there is wide variability in the ASL proficiency of CODAs who do not interpret professionally (CODA Non-Interpreters, $n = 19$), CODAs who are professional interpreters would have a uniformly high level of ASL proficiency (CODA Interpreters, $n = 12$). Under the same logic, we included 15 Non-CODAs with at least 10 years of professional ASL interpreting experience, and 16 Non-CODAs with no interpreting experience at all. Data from 11 subjects who did not meet these criteria were excluded from analysis. The inclusion of these different subject groups also allowed us to additionally test the notion that simultaneous interpreters may have overall increased STM capacity. Data from one additional participant (a Non-CODA Non-Interpreter) were excluded because one condition was omitted due to experimenter error.

All subjects gave consent to be videotaped for data analysis, and were compensated for their participation.

2.1.2. Design

Our main objective was to discover whether using ASL vs. English differentially affected STM span at each stage of the short-term memory process (perception, encoding, recall). Therefore, we manipulated which language was used during each of these stages in a $2 \times 2 \times 2$ within-subjects design. Perception was manipulated by presenting the stimuli in either ASL or in audio-visual English. The internal code used for memory encoding was manipulated by instructing the subjects to shadow the stimulus model while the sequence was being presented by either overtly signing each digit to themselves in ASL, or overtly mouthing each digit in English while remaining silent. Both shadowing tasks require overt and controlled movement of either the hands or the mouth. English shadowing was silent, to avoid introducing differences in acoustic input between the ASL and English encoding conditions. Recall was manipulated by instructing the subjects to report the sequence back from memory in either ASL or English.

To guard against the possibility that reduced span in ASL might simply reflect weaker ASL language skills overall, we included two between-subjects factors: CODA status (CODA vs. Non-CODA) and Interpreter Status (Interpreter vs. Non-Interpreter).

We also included separate conditions to measure articulation rate, basic digit span, and free recall in ASL and English, but those data will not be discussed here. Results from some of these conditions have been reported in Bavelier, Newport, et al. (2008).

2.1.3. Materials

For the digit span stimuli, we filmed a native ASL-English bilingual producing sequences of digits either in ASL or in English. The digits 1–9 were ordered randomly to form lists of increasing length according to the pattern of the WAIS digit span task (Wechsler, 1999), starting at 2 items and proceeding up to 12 items with the constraints that no digit was repeated in a sequence (excepting sequences longer than 9 items until all 9 digits had been used), and that no strings of more than two consecutive digits were allowed (e.g. "3 2 1" would not be acceptable). There were two sequences at each length. The various lists in each language were produced by a native ASL-English bilingual at a rate of one item per second. The signs were produced with a neutral facial expression, and with only those mouth movements that would be natural for native ASL signing. The English items were produced with normal list intonation. Digital video clips were converted to QuickTime format through iMovie software and displayed using either the Matlab computer language (The Math Works Inc., Natick, MA) and the Psychophysical Toolbox routines (Brainard, 1997; Pelli, 1997) (http://psychtoolbox.org) on a Macintosh PowerBook G3 laptop computer (monitor size: 14"), or through Psyscope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh MacBookPro

Table 1
Subject Demographics.

Group	N	Age M (SD)	% Two Deaf parents	% Use ASL at least weekly	Years of interpreting M (SD)	Self-report ASL comprehension: M (SD); scale 1–4	Self-report ASL production: M (SD); scale 1–4
CODA Interpreter	12	35.3 (11.9)	100.0	100.0	14.8 (10.9)	3.68 (.45)	3.50 (.51)
CODA Non-Interpreter	19	24.1 (6.9)	94.7	94.7	0.2 (0.7)	3.26 (.46)	3.00 (.53)
Non-CODA interpreter	15	44.3 (9.1)	0.0	100.0	19.0 (7.6)	2.96 (.46)	2.82 (.52)
Non-CODA Non-interpreter	15	20.1 (0.7)	0.0	93.3	0.0 (0.0)	2.53 (.46)	2.37 (.52)
CODA (Experiment 2 – Letters)	22	21.3 (5.2)	77.3	100.0	1.3 (2.5)	3.38 (.50)	3.19 (.51)

(monitor size = 15"). No participant saw the same sequence more than once.

2.1.4. Procedure

Subjects were tested individually in a 90-min session that ended with the eight digit-span tasks on which we focus here.

Digit span was measured by presenting videos of a person producing a sequence of digits. At the end of each video, the computer paused and the subject recalled the sequence from memory. Following the WAIS procedure, subjects were given two trials at each list length; if they recalled at least one correctly, testing continued. Testing ended when they failed to correctly recall both sequences at a given list length.

At the start of each condition, the experimenter told the subject (in English) whether the sequence would be presented in ASL or English, whether to sign to themselves or silently mouth, and whether to recall the sequence from memory in sign or speech. We describe these eight conditions with the convention that the first letter represents language of presentation (A for ASL, E for English), the second letter represents language of shadowing, and the third letter represents language of recall.

To lighten the cognitive load of the task, we divided the eight conditions into two blocks of four. One block consisted of the conditions that required shadowing in ASL (AAA, AAE, EAA, EAE) and the other block consisted of those that used English memorization (EEE, EEA, AEE, AEA). Within each of these blocks, presentation and recall were factorially manipulated. The order of both blocks and conditions was counterbalanced across subjects. No subject saw the same sequence more than once, but a sequence that was used in the AAA condition for Subject 1 might have been used in the AEE condition for Subject 2. Which list was matched with which condition was balanced within language of presentation across subjects. That is, lists that were filmed in ASL were encountered equally often in AAA, AAE, AEA, and AEE, but never in conditions using English presentation, and vice versa. We acknowledge that, given this limitation, an effect of language presentation could result from differences between the lists using ASL vs. English for presentation (note this issue does not apply to the effect of shadowing or recall language). The fact that Experiment 2 replicates the outcome of Experiment 1 with entirely different stimuli much weakens such an interpretation, however.

2.1.5. Scoring

To measure short-term memory, we consider two measures commonly used with serial recall task: span and score. Span is operationalized as the longest length at which a participant correctly recalls at least one sequence, and may be interpreted as a rough index of capacity: that is, how many items may be held in short-term memory. Score is operationalized as the number of correct trials that a participant has completed before failing both trials at a given length. It is therefore slightly more sensitive, but less straightforward to interpret in terms of memory capacity. We analyzed both spans and scores, but focus our discussion on only those results that were significant in both analyses. Figures show spans; tables include both spans and scores.

Recall rate was scored by measuring the time that participants spent recalling a sequence from memory, from the onset of the first item to the offset of the last item. Onset was identified as the first frame in which the sign's handshape was fully formed and the hand was no longer being raised from a resting position. Offset was identified as either the first frame where the hand was lowered or the handshape began relaxing. From each subject, we measured recall rate from correctly recalled sequences at a list length of four, for all 8 conditions. If a condition had not been videotaped, or if a participant had no correct trials in a given condition, that cell was left blank in computing average recall rate.

ASL fluency was estimated by asking participants to fill in a detailed background language questionnaire which included self-rating on a scale from 1 to 4 on ASL comprehension and production separately, where 1 corresponded to "poorly" and 4 corresponded to "perfectly".

2.2. Results

We analyzed the data using a $(2 \times 2 \times 2) \times (2 \times 2)$ ANOVA with Presentation Language (ASL vs. English), Shadowing Language (ASL vs. English) and Recall Language (ASL vs. English) as within-subjects factors, and CODA

Table 2a
Digit span data – spans.

Condition	Population mean (SD)				Grand mean (SD)
Spans	CODA Interpreter	CODA Non-Interpreter	Non-CODA Interpreter	Non-CODA Non-Interpreter	
AAA	5.67 (1.30)	5.94 (1.18)	6.27 (1.91)	5.67 (1.18)	5.90 (1.40)
AAE	5.33 (0.89)	5.63 (1.07)	6.20 (1.74)	5.20 (1.01)	5.61 (1.26)
AEA	6.25 (1.29)	6.16 (1.30)	6.47 (1.60)	5.93 (1.10)	6.20 (1.31)
AEE	5.83 (1.59)	6.00 (1.05)	6.73 (1.83)	5.87 (0.92)	6.11 (1.38)
EAA	6.58 (1.24)	6.11 (1.15)	7.00 (1.81)	6.27 (1.16)	6.46 (1.37)
EAE	6.42 (1.08)	5.95 (1.09)	6.87 (1.51)	5.87 (0.83)	6.25 (1.19)
EEA	6.25 (1.22)	6.42 (1.02)	7.40 (1.55)	6.00 (1.13)	6.52 (1.31)
EEE	6.08 (1.44)	6.16 (1.26)	6.87 (1.77)	6.80 (1.74)	6.48 (1.56)

Table 2b
Digit Span Data – Scores.

Condition	Population mean (SD)				Grand mean (SD)
Scores	CODA Interpreter	CODA Non-Interpreter	Non-CODA Interpreter	Non-CODA Non-Interpreter	
AAA	8.25 (2.14)	8.74 (1.48)	9.47 (3.60)	8.27 (2.12)	8.70 (2.41)
AAE	7.67 (1.72)	7.89 (1.97)	9.07 (3.37)	7.47 (1.77)	8.03 (2.34)
AEA	9.67 (2.19)	9.37 (2.01)	9.87 (2.70)	9.07 (2.25)	9.48 (2.24)
AEE	8.75 (2.45)	8.89 (2.16)	10.13 (3.14)	8.67 (1.59)	9.11 (2.40)
EAA	10.00 (2.26)	9.21 (1.65)	10.80 (2.98)	9.60 (1.96)	9.85 (2.26)
EAE	10.17 (2.08)	9.11 (1.76)	10.60 (3.00)	8.87 (1.51)	9.62 (2.21)
EEA	9.75 (2.01)	9.79 (1.84)	11.40 (2.87)	9.27 (1.75)	10.05 (2.25)
EEE	9.08 (2.81)	9.58 (2.09)	10.93 (3.20)	10.20 (2.81)	9.97 (2.73)

status (CODA vs. Non-CODA) and Interpreter status (Interpreter vs. Non-Interpreter) as between-subjects factors. We conducted separate analyses with STM spans and STM scores as dependent variables, and report effects that were significant by both measures. See Table 2a (Spans) and Table 2b (Scores) for means and standard deviations for each group.

2.2.1. Within-subjects factors

As shown in Fig. 1A, a main effect of Presentation Language revealed higher digit span when sequences were presented in English than in ASL [span: $F(1, 57) = 22.62$, $p < .001$, $\eta_p^2 = .284$; score: $F(1, 57) = 34.4$, $p < .001$, $\eta_p^2 = .376$].

Fig. 1B shows the main effect of Shadowing Language [span: $F(1, 57) = 12.05$, $p < .002$, $\eta_p^2 = .175$; score: $F(1, 57) = 24.37$, $p < .001$, $\eta_p^2 = .299$]. Digit span was significantly higher when subjects relied on speech-based internal codes (through silent mouthing of English numbers) than when they relied on sign-based codes (through reproducing the sequence on the hands).

The interaction of Presentation and Shadowing was also significant [span: $F(1, 57) = 5.72$, $p < .03$, $\eta_p^2 = .091$; score: $F(1, 57) = 9.90$, $p < .01$, $\eta_p^2 = .148$]. The data, shown in Fig. 2, indicate that English shadowing substantially increased span when presentation was in ASL, but that English shadowing had less of an effect if the sequence had already been presented in English.

Finally, a main effect of Recall indicated that recalling sequences in ASL resulted in *higher* digit span than recalling them in English, as seen in Fig. 1C [span: $F(1, 57) = 5.58$, $p < .03$, $\eta_p^2 = .089$; score: $F(1, 57) = 7.70$, $p < .01$, $\eta_p^2 = .119$].

2.2.2. Recall rate

The ASL advantage[1] at recall was unexpected. Previous research on recall processes suggests that time of articulation is a critical factor (Dosher & Ma, 1998). Therefore, we asked whether the ASL advantage at recall could be attributed to participants performing recall faster in ASL than in English. We conducted a two-way ANOVA with Recall Language as a two-level within-subjects factor (ASL vs. English) and Group as a four-level between-subjects factor (Results do not differ meaningfully if Group is treated as two factors of CODA Status and Interpreter Status). As shown in Table 3, recall was significantly faster in English (2.87 items/s) than in ASL (2.66 items/s; $F(1, 56) = 4.70$, $p < .04$). Neither the main effect of Group [$F(3, 56) = 1.83$, $p = .15$] nor the Group × Recall Language interaction [$F(3, 56) = 1.90$, $p = .14$] was significant. We defer further discussion of the ASL recall advantage until the General Discussion; for now, it is sufficient to note that it cannot be attributed to participants performing recall faster in ASL, because they were in fact faster to recall sequences in English.

No other within-subjects effects reached statistical significance in spans and scores.

2.2.3. Proficiency ratings

When performing research with bilinguals, it is important to take into account subjects' relative proficiency in their two languages. By virtue of being native English speakers, the groups under investigation are unlikely to vary much in their English proficiency, and even the most

[1] The ASL advantage at recall may also help to explain why sign language interpreters often report voice-to-sign interpreting to be easier than sign-to-voice. We thank Bencie Woll for offering this observation (personal communication, April 11, 2005).

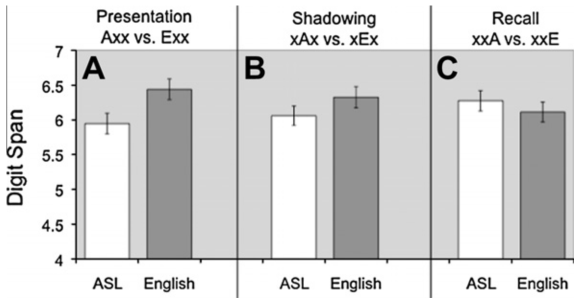

Fig. 1. Using speech for presentation and shadowing increases span, but recall favors sign. Error bars represent SEM.

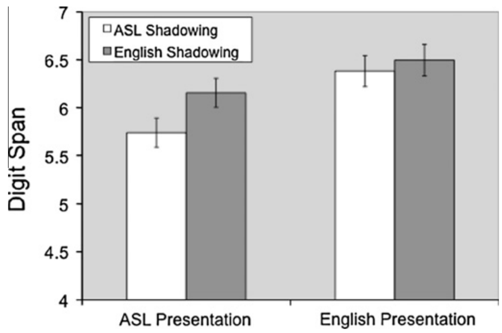

Fig. 2. English shadowing is effective at increasing span when presentation is in ASL, but makes less of a difference when presentation is already in speech. Error bars represent SEM.

proficient signers are likely to be English dominant. In contrast, we *do* expect the groups to vary significantly in their ASL proficiency, and therefore in how English dominant they may be. For a quantitative test of proficiency, we entered participants' self-ratings for ASL comprehension and production into a 2×4 ANOVA, with Mode (Comprehension & Production) as a within-subjects factor and Group (CODA Interpreter, CODA Non-Interpreter, Non-CODA Interpreter, and Non-CODA Non-Interpreter) as a between-subjects factor. (Results do not differ meaningfully if Group is treated as two factors of CODA Status and Interpreter Status). Mean proficiency ratings are shown in Table 1. The main effect of Group was significant [$F(3, 57) = 14.68$, $p < .001$, $\eta_p^2 = .436$]. A Tukey-HSD test confirmed that, as expected, the CODA Interpreters rated themselves significantly more highly than all other groups, and the Non-CODA Non-Interpreters rated themselves as less proficient than all other groups. CODA Non-Interpreters and Non-CODA Interpreters fell in between and did not differ from each other. Participants consistently rated their comprehension higher than their production, as indicated by a main effect of Mode [$F(1, 57) = 11.08$, $p < .003$, $\eta_p^2 = .163$]. The Group x Mode interaction was not significant ($F < 1$).

2.2.4. Between-subjects factors

Having established that our participant groups do differ in their ASL proficiency, we test whether lower proficiency in ASL might account for the disadvantages observed when presentation and shadowing occur in ASL. If so, we would expect these effects to be exaggerated in subjects of lower ASL proficiency (e.g. Non-CODAs and Non-interpreters). To address this issue, we ask whether CODA status or Interpreter status influence overall memory capacity, and/or interact with the within-subjects factors of Presentation Language, Shadowing Language, or Recall Language. We look at this issue by returning to the omnibus analysis with STM spans and scores as dependent variables.

There was no main effect of CODA status [spans: $F(1, 57) = 1.06$, $p = .31$, $\eta_p^2 = .018$; scores: $F(1, 57) = .94$, $p = .34$, $\eta_p^2 = .016$]. CODA status did not interact with any within-subject factors (all Fs < 1.7). Thus, subjects who acquired ASL as a second language showed similar patterns of perception, encoding, and recall as subjects for who acquired ASL from birth. However, age of acquisition is only one measure of proficiency.

Some late learners attain a high degree of mastery, while some CODAs experience first language attrition. Highly proficient late learners are likely to be interpreters, whereas CODAs who have lost some ASL skill are unlikely to be interpreters; thus, comparing interpreters to Non-interpreters is one way of capturing these differences. Spoken language interpreters have been shown to have increased capacity in some, linguistic memory tasks, whether as a result of training or by virtue of self-selection (Christoffels, de Groot, & Kroll, 2006); however, we know of no comparable studies of sign language interpreters. The results below provide a first insight in this question.

There was no main effect of Interpreter status [spans: $F(1, 57) = 1.95$, $p = .17$, $\eta_p^2 = .033$; scores: $F(1, 57) = .94$, $p = .38$, $\eta_p^2 = .016$]. Despite a trend towards increased STM capacity in interpreters (span = 6.39, score = 9.73) over Non-interpreters (span = 6.00, score = 9.00), the difference was not reliable. Interpreter status did not interact significantly with any within-subject factors (all Fs < 3.1).

No other between-subjects effects reached statistical significance in spans and scores.

2.3. Discussion

The present study is the first to compare sign and speech during different processing stages of STM,

Table 3
Recall rate (items/s).

Recall language	Population mean recall rate: items/s (SD)				Grand mean (SD)
	CODA interpreter	CODA Non-interpreter	Non-CODA interpreter	Non-CODA Non-Interpreter	
ASL	3.05 (.96)	2.73 (.74)	2.71 (.93)	2.20 (1.03)	2.66 (.95)
English	3.11 (.90)	2.91 (.94)	2.69 (.99)	2.79 (.74)	2.87 (.91)

providing new insight into both commonalities and differences between the two language modalities. The results show that the stages of STM are differentially affected by the language used. English during perception and encoding leads to higher serial STM span than ASL; the reverse pattern is observed during recall. Before discussing the implications of these results, it is important to address two possible alternative explanations.

The first one is the possibility that these results are a byproduct of the fact that for our participants, ASL is a weaker language. The significant differences in participants' self-rated proficiency enabled us to test the predictions of proficiency-based explanations for our data. If ASL proficiency were responsible for the pattern of results in E1, we would expect less-proficient participants to be worse overall. The analyses show that there was no main effect of CODA status or Interpreter status on the STM spans or scores measured. Furthermore, under a fluency explanation, we would expect all STM stages to be similarly affected; we know of no *a priori* reason to suspect that some stages of the STM process should be more affected by language proficiency than others. However, we found that presentation and encoding patterned differently than recall. Even harder to reconcile with an ASL proficiency explanation is the finding that using ASL during recall significantly *increased* performance, while it decreased performance for presentation and shadowing. Finally, under the ASL proficiency account, the disadvantages we observed for ASL Presentation and Shadowing should be exaggerated in subjects of lower proficiency. However, no such interactions were noted in the omnibus analysis. We tested this possibility more directly by comparing the most and least ASL proficient participants (CODA Interpreter vs. Non-CODA Non-Interpreter). In so doing, we found no trace of a Group × Presentation interaction [spans: $F(1, 25) = 0.00$, $p = .99$, $\eta_p^2 = .00$; scores: $F(1, 25) = .01$, $p = .93$, $\eta_p^2 = .00$], nor is there a Group x Shadowing interaction [$F(1, 25) = 1.44$, $p = .24$, $\eta_p^2 = .054$; scores: $F(1, 25) = 1.45$, $p = .24$, $\eta_p^2 = .055$]. Thus, although our participants did differ in their ASL fluency, such differences do not satisfactorily explain the pattern of results that we found. We turn now to a second potential confound: phonological similarity among ASL digits.

It is possible that the results of Experiment 1 are a byproduct of the fact that digits are more phonologically similar in ASL than in English. The nature of the stimuli is always a concern in cross-linguistic studies (Bavelier et al., 2006; Wilson & Emmorey, 2006). It could be argued that the use of digits led to greater phonological similarity in sign than in speech. Indeed, fingerspelled digits are phonologically more similar to each other than their spoken equivalents. This could have resulted in a lower span when using sign-based representations. However, several factors render this explanation unlikely. First, signers have equivalent spans whether phonologically similar digits are being used or phonologically dissimilar letters are being used (see Bavelier, Newport, et al., 2008, Experiments 1 and 3). Importantly, such fingerspelled materials are not especially challenging for signers. These items are short, easy to articulate, and accordingly lead to a span size of about 5 ± 1, which is among the longest span sizes that have been elicited in signers (Bavelier, Newport, et al., 2008; Hall & Bavelier, 2010). Second, the finding of a greater span when recall is in ASL rather than in English is not consistent with a phonological similarity interpretation of the results. Indeed, phonological similarity at recall should be detrimental to performance; instead, recall in ASL led to greater span, if anything. Nevertheless, we acknowledge that the ideal stimuli would be equally phonological simple and dissimilar in both ASL and English, such as the letters used by Bavelier et al. (2006). To address this concern, we conducted Experiment 2.

3. Experiment 2

3.1. Method

3.1.1. Subjects

We tested 28 additional ASL-English bilinguals, all of whom had at least one Deaf parent, and had been exposed to ASL and English since birth. Based on the results of Experiment 1, we included subjects with varying levels of interpreting experience and did not consider this factor in our analysis. Due to experimenter error, one subject did not perform all eight conditions and thus was excluded from analysis. Either equipment or experimenter error resulted in six of the remaining subjects being prematurely stopped in at least one condition. To guard against the possibility that these subjects' data could inflate any effects we might find, we excluded all but the one subject who should have continued in the EEA condition, since if anything, her data should work against our hypothesis. Demographic information for the final 22 subjects in Experiment 2 is included in Table 1.

3.1.2. Design

The design was identical to Experiment 1, except that there was only one group and therefore no between-subjects factors.

3.1.3. Materials

Experiment 2 used a carefully selected set of letters (see Bavelier et al., 2006) to test the possibility that the ASL disadvantages for Presentation and Shadowing in Experiment 1 were a result of phonological similarity within the ASL digits. The challenge in selecting a set of 9 letters lies in minimizing their phonological similarity in English (e.g. /bi/, /si/ /di/, /pi/, /ti/, etc.). After discarding vowels (to prevent subjects from chunking word-like sequences), we selected G, H, K, L, M, R, S, and Y as the most phonologically distinct letter names. For the 9th letter, we added P on the grounds that it was similar to G in English and also to K in ASL. However, P and K did not appear together until length 8, and P and G did not appear together until length 9, by which point almost all subjects had reached their span. We consulted 3 native ASL signers as informants (two Deaf, one hearing), who agreed that, with the exception of P/K, these letters are phonologically dissimilar in ASL. We are therefore confident that if the same pattern of results emerges in Experiment 2, it cannot be attributed to phonological similarity within the items for either ASL or English.

We were also careful to avoid sequences of more than two consecutive letters of the alphabet (e.g. KLM), as well as sequences that formed sentence-like chunks (e.g. YR) or acronyms (e.g. PMS). It is also worth acknowledging that, like the digits 1–9, these ASL letters lack movement, which is a property of lexical signs in ASL. It is thus possible that neither digits nor letters are ideal for testing short-term memory in signers; however, given that digits or letter spans are the standard in clinical and educational settings, we believe these items are important to study.

3.1.4. Procedure

The procedure was identical to Experiment 1.

3.2. Results

We conducted a $2 \times 2 \times 2$ ANOVA with Presentation Language (ASL vs. English), Shadowing Language (ASL vs. English), and Recall Language (ASL vs. English) as within-subjects factors (Table 4). Despite having roughly one-third the sample size of Experiment 1, the pattern of data was strikingly similar. In the span analysis, we found an advantage for English during presentation [$F(1, 21) = 4.56$, $p < .05$, $\eta_p^2 = .18$; Fig. 3A] and during shadowing [$F(1, 21) = 7.82$, $p < .02$, $\eta_p^2 = .27$; Fig. 3B], and a trend toward an ASL advantage at recall ($F(1, 21) = 3.28$, $p = .085$, $\eta_p^2 = .14$; Fig. 3C). In addition, there was a trend in the predicted direction for a Presentation × Shadowing interaction [$F(1, 21) = 4.19$, $p = .053$, $\eta_p^2 = .17$; Fig. 4]. Note that these reported p-values are two-tailed, whereas our hypothesis makes explicitly one-tailed predictions. In the scores analysis, only two of the above effects reached significance: the main effect of Shadowing Language [$F(1, 21) = 7.62$, $p < .02$, $\eta_p^2 = .27$] and the Presentation × Shadowing interaction [$F(1, 21) = 5.82$, $p < .03$, $\eta_p^2 = .22$]. The scores analysis also yielded a Presentation × Shadowing x Recall interaction [$F(1, 21) = 6.13$, $p < .03$, $\eta_p^2 = .23$] that was not significant in the spans analysis [$F(1, 21) = 2.12$, $p = .15$, $\eta_p^2 = .095$], and thus not interpreted further.

Thus, the results of Experiment 2 generally mirror those of Experiment 1 with effect sizes that are at least equal to - and in some cases larger than- those in Experiment 1.

3.3. Discussion

The goal of Experiment 2 was to test the viability of an alternative hypothesis by which the ASL disadvantage seen in Experiment 1 for presentation and shadowing might simply be an artifact of ASL digits being phonologically similar. Under that hypothesis, those two effects should have disappeared in Experiment 2, which used letters that are phonologically dissimilar in both ASL and English. It appears that the main effect of Presentation did become weaker in Experiment 2, suggesting that phonological similarity might be important in early stages of STM. However, the fact that the main effect of Shadowing and the Presentation × Shadowing interaction persisted with these new materials and a smaller sample strongly suggests that these effects are not due to phonological similarity. Rather Experiment 2 provides further support for the hypothesis that using speech-based encoding (through shadowing or as a byproduct of spoken presentation) leads to higher serial STM span than sign. Finally, Experiment 2 replicates the observation that the serial span discrepancy arises prior to recall, where using sign tends to lead to a higher span than speech.

Table 4
Letter Span Data – Spans & Scores.

Condition	Span (SD)	Score (SD)
AAA	5.05 (.84)	7.14 (1.46)
AAE	4.86 (.98)	7.05 (1.59)
AEA	5.86 (1.13)	8.55 (1.89)
AEE	5.27 (.88)	7.64 (1.46)
EAA	5.50 (1.10)	7.91 (1.85)
EAE	5.23 (.92)	7.55 (1.53)
EEA	5.64 (1.46)	7.82 (2.13)
EEE	5.59 (1.10)	8.09 (1.95)

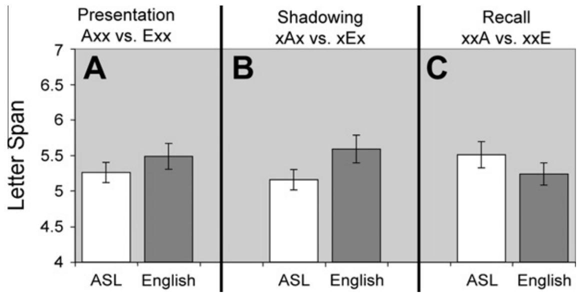

Fig. 3. Using speech for presentation and shadowing increases span, but recall favors sign, even when phonologically dissimilar stimuli (letters) are used. Error bars represent SEM.

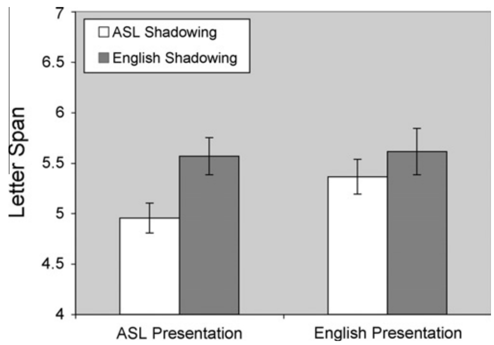

Fig. 4. English shadowing increases span when presentation is in ASL, but makes less of a difference when presentation is already in speech. Error bars represent SEM.

4. General discussion

In two experiments testing serial short-term memory for linguistic materials, we showed that perception, encoding, and recall are differentially affected by the use of speech vs. sign. The use of English during presentation and/or during shadowing leads to higher STM span, whereas the use of ASL during recall leads to higher STM span. In addition, shadowing in English was most effective in raising the span when presentation had been in ASL.

The advantage for speech during perception (Figs. 1A and 3A) is in accord with previous research showing STM advantages for spoken over signed stimuli (Bavelier, Newman, et al., 2008; Bavelier, Newport, et al., 2008; Boutla et al. 2004; Hamilton & Holzman, 1989; Hoemann & Blama, 1992; Krakow & Hanson, 1985). However, these studies have been confounded by factors such as hearing status, individual differences, and other nuisance variables such as item duration and frequency. Most problematically, they have confounded presentation modality with subjects' internal codes. The present findings establish that presentation in English yields better serial recall than presentation in ASL, even when controlling for the codes used during memorization and recall, at least for hearing ASL-English bilinguals; whether spoken presentation is also advantageous to deaf signers is uncertain at best. In addition, this result was established using stimuli that are part of a dynamic language system in both languages, rather than written representations of language.

The most robust effect across both experiments was the advantage for encoding in English. This indicates that subjects' internal codes contribute to the serial span discrepancy, independent of the perceptual features of the input. If the speech advantage were purely a byproduct of peripheral acoustic processes, then silently recoding a sequence presented in ASL into speech-based representations should not have resulted in increased span. Yet, this is exactly what was found (see Figs. 1B and 3B). These results are consistent with earlier studies that have also documented possible advantages for speech-based representations in short-term memory, although their interpretation is complicated by the lack of direct manipulation (e.g. Conrad, 1970; Hamilton & Holzman, 1989; Hanson, 1982; Koo et al., 2008; Krakow & Hanson, 1985; Lichtenstein, 1998; Miller, 2007).

The joint contributions of perceptual and encoding factors are illustrated by the interaction between presentation language and shadowing language. Shadowing in English was most effective at increasing span when presentation was in ASL (Figs. 2 and 4). This pattern of interaction sheds light on two main properties of short-term memory processes. First, using speech-based codes in memory facilitates verbal short-term memory for serial lists. So much so, that the span in the AEA condition was numerically higher than that in the AAA condition even though AEA requires two translations along the way. Such translations would be expected to result in lesser span due to their memory taxing demands. Second, this pattern of interaction follows naturally from the proposal that spoken sequences gain automatic access to the speech-based phonological loop (Baddeley, 1986). Mouthing silently is another route to speech-based representations (Campbell, 1992; Crowder, 1983). Indeed, these data strongly resemble those of Hamilton and Holzman (1989), who compared speech-only, sign-only, and bimodal presentation in various populations. They found that for hearing signers, adding sign to speech made little difference, but adding speech to sign improved memory. For their deaf signers (whose speech skills were variable), bimodal presentation yielded better memory than either speech or sign alone. Thus, it is possible that the most critical component of the language of presentation advantage may not be its acoustic characteristics (which could still play a role in hearing subjects), but rather the engagement of speech-based memory representations. Taken together, these results suggest that although sign and speech share the same memory architecture, they may differ in how efficiently linguistic representations move through the phonological loop. Specifically, rehearsal and chunking processes, which appear to be supplemental strategies that increase basic span size (Cowan, 2001; Cowan et al., 2005) may be given less weight when dealing with sign-based representations. Because serial recall tasks are designed to focus on the phonological loop to the exclusion of other working memory buffers, they highlight these differences. Why rehearsal and chunking may be less dominant strategies when it comes to sign remains to be elucidated.

One surprising result that emerged from the data was a subtle advantage for ASL during recall (Figs. 1C and 3C). This result argues against any global disadvantage for sign languages in memory tasks. More specifically, the data strongly refute the notion that the shorter span observed for sign languages arises during recall, whether as a function of slower motor output systems (Klima & Bellugi, 1979; Lichtenstein, 1998; Marschark, 1996; Marschark & Mayer, 1998), or other reasons. Most previous studies have used written recall in an effort to equate performance between deaf and hearing groups, but this may have unduly hindered signing subjects. In the present study, performing recall in sign may have allowed participants to retain speech-based representations in memory while performing recall in a modality that did not interfere with those representations. That is, when producing spoken recall, the output feeds back to the same input buffer, potentially causing interference. Monitoring for sign language, however, may not be subject to such interference. Recent work by Emmorey, Bosworth, and Kraljic (2009) suggests that sign monitoring may be kinesthetic, whereas sign perception is visual. Consistent with our findings, this account would then predict less interference between sign perception and sign monitoring. Alternatively, a small literature is emerging that documents memory advantages for stimuli associated with manual gestures (Cohen & Otterbein, 1992; Feyereisen, 2006). Whether the present result reflects such a process remains to be addressed.

The present results offer new support for an old theory: namely, that although signed input can be processed by a sign-based phonological loop (akin to the phonological loop described for speakers), some still-unspecified property of speech-based representations renders them more efficient than sign-based representations for the purposes of serial recall tasks (Conrad, 1970; Hamilton & Holzman, 1989; Hanson, 1982; Koo et al., 2008; Krakow & Hanson, 1985; Lichtenstein, 1998; Miller, 2007). Several aspects of the current data support this view. Serial recall performance is best when speech-based memory representations are engaged, either through automatic activation following English presentation, or through silent speech-based shadowing following ASL presentation. Because the data come from a within-subjects design with hearing participants, the differences cannot be attributed to impoverished mnemonic abilities in the Deaf, lack of full access to speech-based representations, or between-subject variability. The disadvantage for sign is also not simply a byproduct of testing participants with weaker language skills in ASL. If this had been the case, we would have expected to see exaggerated English advantages for those with weaker sign skills, but this was not the case. Crucially, our study sheds new light on the original theory by establishing that sign-based representations are not systematically counterproductive in serial STM tasks. Rather, use of sign led to shorter span only during perception and encoding, but not during recall. If sign use systematically led to weaker serial STM, we would have expected a disadvantage for sign to appear in all three stages, whereas the data show that using ASL during recall tended to increase span.

Although we have occasionally used terminology such as sign-based representations "reducing" memory, we wish to note that when considered in the context of other human memory systems, STM in sign languages is unexceptional. Cowan (2001) documents that almost all types of memory show a capacity of around 4–5 items, except for verbal memory in speakers. Seen in this light, the puzzle is not why sign-based memory may be around 4–5 items, but rather why speech-based memory is peculiarly high.

In thinking about possible differences between a sign-based and speech-based phonological loop, three main possibilities come to mind: (1) phonological weight, (2) the phonological store where phonological traces contact long-term memory representations, and (3) the articulatory loop or rehearsal process that allows phonological traces to be maintained in short-term memory.

Phonological weight has recently been suggested as a potential source of the serial span discrepancy by Gozzi, Geraci, Cecchetto, Perugini, and Papagno (in press). Following Brentari (1998), they take phonological weight to be influenced by number of hands, amount and type of movement, and presence/absence of handshape change within the sign. Gozzi et al., like Geraci et al. (2008), used phonologically light lexical signs, but still found a difference in STM between spoken Italian and LIS. Consequently, they suggest that even light lexical signs may be phonologically heavier than some spoken words. Future studies manipulating phonological weight would be helpful in evaluating this account; however, it is important to note that our digit and letter materials are the lightest possible signed elements, involving only one hand, no movement, and no internal handshape change. The fact that we still observe a serial span discrepancy with such items suggests that either phonological weight is an additional independent factor, or that even the lightest ASL signs are heavier than English digit and letter names, a position for which there is at present little support.

We now consider whether the difference we observe might arise in the phonological store or the rehearsal loop. While the present study cannot differentiate between these two processes, other data in the field point to differences in the rehearsal process when handling sign vs. speech. In a meta-analysis of 27 short-term memory studies of signers and speakers, Hall and Bavelier (2010) found that articulation rate strongly predicted STM span for speech ($r^2 = .74$). This finding is in line with the work of Baddeley and others proposing that the amount of information rehearsed at any one time constrains span size (Baddeley, Liews, & Vallar, 1984; see also Mueller & Krawitz, 2009). For signers, there was no correlation between articulation duration of the items and span size for the same items ($r^2 = .00$), suggesting a lesser reliance on rehearsal. Additionally, a brain imaging study comparing deaf signers and hearing speakers while performing a serial STM task found a marked decrease in activation during the rehearsal stage of STM in deaf signers, in accord with the view that this stage of processing is less involved in signers. In contrast, much stronger activation in deaf signers was found during recall, suggesting a different trade-off between covert rehearsal and overt response in signers and speakers (Bavelier, Newman, et al., 2008). This work combined with that of others support the proposal that signers may rely less predominantly on rehearsal for serial STM maintenance than speakers.

In this view, speakers may be highly specialized for rote rehearsal during verbal memory, whereas signers may be more likely to invoke several internal codes for linguistic memory. These would include sign-based phonological coding as illustrated by the work of Wilson and Emmorey (1997), Wilson and Emmorey (1998, Wilson and Emmorey (2003), but also a variety of other types of memory codes. For example, it is well-known that in some tasks, even hearing speakers rely on conceptual and spatial memory codes (McElree, Foraker, & Dyer, 2003; Postle, D'Esposito, & Corkin, 2005; Potter, 1999; Potter, Representational Buffers: The Eye-Mind Hypothesis in Picture Perception, Visual Search, & Language Processes., 1983). The proposal that multiple codes conspire to support short-term memory processes is far from new (Paivio, 1986; Paivio, 2007), and thus the proposal that the very rich information afforded by sign language may lead to multiple coding in short-term memory is unsurprising (see the Ease of Language Understanding model by Ronnberg, Rudner, Foo & Lunner, 2008 for a similar view). In this view, signers may make full use of the redundancy that the memory system has to offer during short-term memory task. The more surprising feature may actually be the heavy reliance of speakers on one code during linguistic serial STM span. Although speakers tested with other stimuli or under other task requirements also exhibit a multiple coding strategy (Postle & Hamidid, 2007; Postle, Idzikowski, Sala, Logie, & Baddeley, 2006; Wickens, Nield, Tuber, & Wickens, 1973), it is often the case that a speech-coding strategy dominates in hearing population. This effect is so strong that contamination by a speech-coding strategy is a constant concern in studies of visuo-spatial memory in speakers. It is possible that this high degree of specialization in speakers may be due to an intrinsic advantage of a speech-code when verbal material has to be serially recalled. Alternatively, this high degree of specialization may rather result from training, in particular from the acquisition of explicit phonetic maintenance and analysis when learning to read. The study of non-literate but otherwise typical adults should be able to shed much needed light on this issue.

Clarifying the causal links between STM span and other cognitive processes is of utmost importance for education and testing. Many standardized IQ tests and educational placement assessments include serial recall tasks such as the digit span. For example, STM span may be a good predictor of reading skill in hearing populations, since hearing readers have a strong tendency to rely primarily on speech-based codes for both types of tasks. However, the evidence just reviewed indicates that Deaf subjects rely on a variety of different codes for STM. Performance on serial tasks may therefore have different implications for reading in hearing speakers and Deaf signers. The experiments presented here suggest that intensive training in the use of speech-based codes might indeed increase STM span, but there is no evidence demonstrating that this would be accompanied by a commensurate increase in reading or language skill. It should be fruitful in future research to explore how the various internal codes Deaf individuals naturally rely on may influence reading, rather than simply focusing on those that are most prevalent among hearing individuals.

Acknowledgments

This work was supported in part by the Charles A. Dana Foundation and by NIH grants DC04418 to Daphne Bavelier and HD051030 to Victor Ferreira. In addition, the authors would like to thank Mrim Boutla, Elissa Newport, Ted Supalla, Matt Dye, Peter Hauser, Dara Baril, and Raylene Paludneviciene for helpful discussions, and Nicole Schaeffer, Darshana Tuladhar, Jason Listman, Jhona Smith, and Brooke Schmeltz for assistance with data collection and coding. We are also grateful to the Psychology Departments at RIT and Gallaudet University as well as the staff at Camp Mark 7 for allowing us to use their space for testing.

Appendix A

See Figs. A1 and A2.

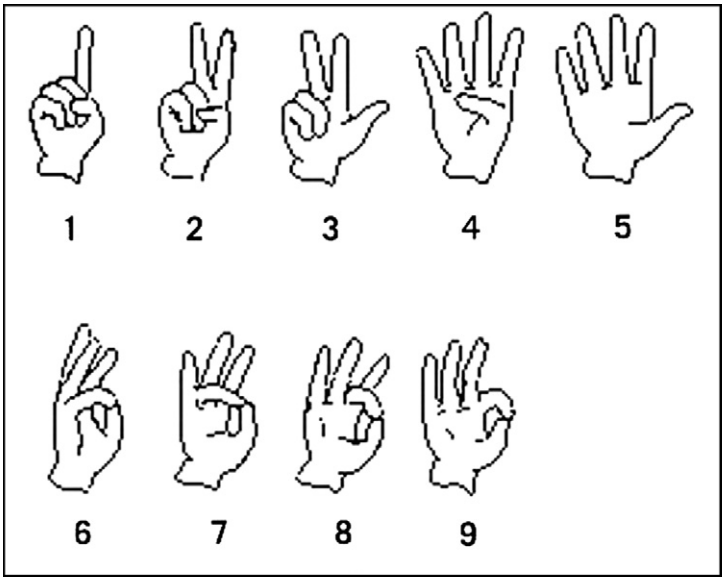

Fig. A1. Illustrations of the ASL digit handshapes used in Experiment 1. Note that participants in the study viewed videos of a native signer signing each list.

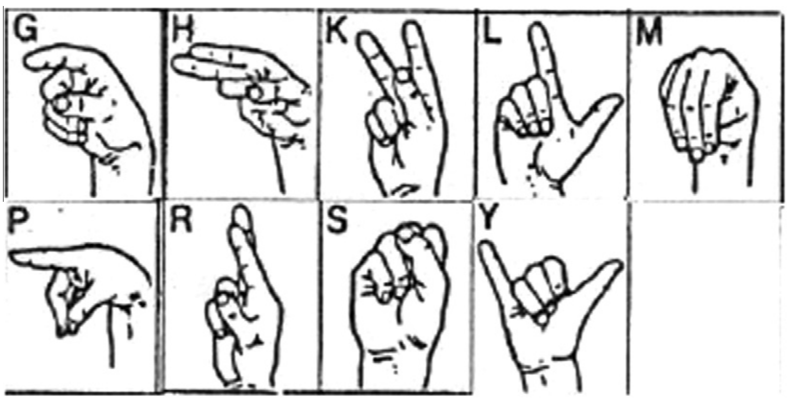

Fig. A2. Illustrations of the ASL letter handshapes used in Experiment 2.

References

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

Baddeley, A., Liews, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36*, 233–252.

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575–589.

Bavelier, D., Newman, A. J., Mukherjee, M., Hauser, P., Kemeny, S., Braun, A., et al. (2008). Encoding, rehearsal, and recall in signers and speakers: Shared network but differential engagement. *Cerebral Cortex, 18*(10), 2263–2274.

Bavelier, D., Newport, E. L., Hall, M. L., Supalla, T., & Boutla, M. (2006). Persistent difference in short-term memory between sign and speech: Implications for cross-linguistic comparisons. *Psychological Science, 17*(2), 1090–1092.

Bavelier, D., Newport, E. L., Hall, M. L., Supalla, T., & Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: Implications for models of short-term memory. *Cognition, 107*(2), 433–459.

Bellugi, U., & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition, 1*, 173–200.

Bellugi, U., Klima, E. S., & Siple, P. (1975). Remembering in signs. *Cognition, 3*(2), 93–125.

Boutla, M., Supalla, T., Newport, E. L., & Bavelier, D. (2004). Short-term memory span: Insights from sign language. *Nature Neuroscience, 7*(9), 997–1002.

Brainard, D. G. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Campbell, R. (1992). Speech in the head? Rhyme skill, reading, and immediate memory in the Deaf. In D. Reisberg (Ed.), *Auditory Imagery* (pp. 73–93). Hillsdale, NJ: L. Erlbaum Associates.

Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology, 45A*(2), 177–192.

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. K. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language, 54*, 324–345.

Cohen, R. L., & Otterbein, N. (1992). The mnemonic effect of speech gestures: Pantomimic and non-pantomimic gestures compared. *European Journal of Cognitive Psychology, 4*(2), 113–139.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257–271.

Conrad, R. (1970). Short-term memory processes in the deaf. *British Journal of Psychology, 61*(2), 179–195.

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

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science, 18*, 275–279.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100.

Crowder, R. G. (1983). The purity of auditory memory. *Philosophical Transactions of the Royal Society of London, B, 302*, 251–265.

Dosher, B. A., & Ma, J.-J. (1998). Output loss or rehearsal loop? Output-time vs. pronunciation-time limits in immediate recall for forgetting-matched materials. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*(2), 316–335.

Elliot, J. M. (1992). Forward digit span and articulation speed for Malay, English and two Chinese dialects. *Perceptual and Motor Skills, 74*, 294–295.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*, 43–51.

Emmorey, K., Bosworth, R., & Kraljic, T. (2009). Visual feedback and self-monitoring of sign language. *Journal of Memory and Language, 61*(3), 398–411.

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology, 18*, 185–205.

Friedman, L. (1976). *Phonology of a soundless language: phonological structure of the American Sign Language*. Unpublished doctoral dissertation, University of California, Berkeley.

Furth, H. G. (1971). Linguistic deficiency and thinking: Research with deaf subjects 1964–1969. *Psychological Bulletin, 76*(1), 58–72.

Geraci, C., Gozzi, M., Papagno, C., & Cecchetto, C. (2008). How grammar can cope with limited short-term memory: Simultaneity and seriality in sign languages. *Cognition, 106*, 780–804.

Gozzi, M., Geraci, C., Cecchetto, C., Perugini, M., Papagno, C. (in press). Looking for an explanation for the low sign span. Is order involved? *Journal of Deaf Studies and Deaf Education*, Advanced Access published August 2, 2010, doi: 10.1093/deafed/enq035.

Hall, M., & Bavelier, D. (2010). Working memory, deafness, and sign language. In M. Marschark & P. A. Spencer (Eds.). *Oxford handbook of deaf studies, language, and education* (Vol. 2, pp. 458–472). London: Oxford University Press.

Hamilton, H., & Holzman, T. G. (1989). Linguistic encoding in short-term memory as a function of stimulus type. *Memory & Cognition, 17*(5), 541–550.

Hanson, V. L. (1982). Short-term recall by Deaf signers of American Sign Language: Implications of encoding strategy for order recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*(6), 572–583.

Hanson, V. L., & Lichtenstein, E. H. (1990). Short-term memory coding by Deaf signers: The primary language coding hypothesis reconsidered. *Cognitive Psychology, 22*, 211–224.

Hoemann, H. W., & Blama, L. S. (1992). Memory span for signs & digits forward & backward in beginning level ASL students. *Sign Language Studies, 76*, 277–284.

Klima, E. S., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Koo, D., Crain, K., LaSasso, C., & Eden, G. F. (2008). Phonological awareness and short-term memory in hearing and deaf individuals of different communication backgrounds. *Annals of the New York Academy of Sciences, 1145*, 83–99.

Krakow, R. A., & Hanson, V. L. (1985). Deaf signers and serial recall in the visual modality: Memory for signs, fingerspelling, and print. *Memory & Cognition, 13*(3), 265–272.

Liben, L. S., & Drury, A. M. (1977). Short-term memory in deaf and hearing children in relation to stimulus characteristics. *Journal of Experimental Child Psychology, 24*, 60–73.

Lichtenstein, E. H. (1998). The relationships between reading processes and English skills of Deaf college students. *Journal of Deaf Studies and Deaf Education, 3*(2), 80–134.

Logan, K., Maybery, M., & Fletcher, J. (1996). The short-term memory of profoundly deaf people for words, signs, and abstract spatial stimuli. *Applied Cognitive Psychology, 10*, 105–119.

MacSweeney, M., Campbell, R., & Donlan, C. (1996). Varieties of short-term memory coding in Deaf teenagers. *Journal of Deaf Studies and Deaf Education, 1*(4), 249–262.

Marschark, M. (1996, November). Influences of signed and spoken language on memory span. Paper presented at the 37th Annual Meeting of the Psychonomic Society, Chicago.

Marschark, M., & Mayer, T. S. (1998). Interactions of language and memory in deaf children and adults. *Scandinavian Journal of Psychology, 39*, 145–148.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48*(1), 67–91.

Miller, P. (2007). The role of spoken and sign languages in the retention of written words by prelingually deafened native signers. *Journal of Deaf Studies and Deaf Education, 12*(2), 184–208.

Mueller, S. T., & Krawitz, A. (2009). Reconsidering the two-second decay hypothesis in verbal working memory. *Journal of Mathematical Psychology, 53*, 14–25.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum Associates.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*, 398–422.

Pintner, R., & Paterson, D. G. (1917). A comparison of deaf and hearing children in visual memory for digits. *Journal of Experimental Psychology, 2*, 76–88.

Poizner, H., & Tallal, P. (1987). Temporal processing in deaf signers. *Brain and Language, 30*(1), 52–62.

Postle, B. R., D'Esposito, M., & Corkin, S. (2005). Effects of verbal and nonverbal interference on spatial and object visual working memory. *Memory and Cognition, 33*(2), 203–212.

Postle, B. R., & Hamidid, M. (2007). Nonvisual codes and nonvisual brain areas support visual working memory. *Cerebral Cortex, 17*(9), 2151–2162.

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology, 59*(1), 100–120.

Potter, M.C. (1983). *Representational Buffers: The Eye-Mind Hypothesis in Picture Perception, Reading, and Visual Search*, in *Eye Movements in Reading: Perceptual and Language Processes*.

Potter, M.C. (1999). *Understanding sentences and scenes: The role of conceptual short-term memory*, in *Fleeting Memories*, V. Coltheart, (Ed.), MIT Press: Cambridge.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47*(Suppl. 2), S171–S177.

Rönnberg, J., Rudner, M., & Ingvar, M. (2004). Neural correlates of working memory for sign language. *Cognitive Brain Research, 20*, 165–182.

Rudner, M., Andin, J., & Rönnberg, J. (2009). Working memory, deafness and sign language. *Scandinavian Journal of Psychology, 50*, 495–505.

Shand, M. A. (1982). Sign-based short-term coding of American Sign Language signs and printed English words by congenitally deaf signers. *Cognitive Psychology, 14*, 1–12.

Wallace, G., & Corballis, M. C. (1973). Short-term memory and coding strategies in the deaf. *Journal of Experimental Psychology, 99*(3), 334–348.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Harcourt Brace and Company.

Wickens, D. D., Nield, A. F., Tuber, D. S., & Wickens, C. D. (1973). Stimulus selection as a function of CS1–CS2 internal in compound classical conditioning of cats. *Journal of Comparative & Physiological Psychology, 85*(2), 295–303.

Wilson, M., & Emmorey, K. (1997). A visuospatial "phonological loop" in working memory: Evidence from American Sign Language. *Memory & Cognition, 25*(3), 313–320.

Wilson, M., & Emmorey, K. (1998). A "word length effect" for sign language: Further evidence for the role of sign language in structuring working memory. *Memory & Cognition, 26*(3), 584–590.

Wilson, M., & Emmorey, K. (2003). The effect of irrelevant visual input on working memory for sign language. *Journal of Deaf Studies and Deaf Education, 8*(2), 97–103.

Wilson, M., & Emmorey, K. (2006). Comparing sign language and speech reveals a universal limit on short-term memory capacity. *Psychological Science, 17*, 682–683.